

METHODOLOGY

Open Access



# A primer on continuous-time modeling in educational research: an exemplary application of a continuous-time latent curve model with structured residuals (CT-LCM-SR) to PISA Data

Julian F. Lohmann<sup>1,2\*</sup> , Steffen Zitzmann<sup>2</sup>, Manuel C. Voelkle<sup>3</sup> and Martin Hecht<sup>2</sup> 

\*Correspondence:  
jlohm@ipl.uni-kiel.de

<sup>1</sup> Institute for Psychology of Learning and Instruction, Kiel University, Olshausenstrasse 75, 24118 Kiel, Germany

<sup>2</sup> Hector Research Institute of Education Sciences and Psychology, University of Tübingen, Tübingen, Germany

<sup>3</sup> Department of Psychology, Humboldt University Berlin, Berlin, Germany

## Abstract

One major challenge of longitudinal data analysis is to find an appropriate statistical model that corresponds to the theory of change and the research questions at hand. In the present article, we argue that *continuous-time models* are well suited to study the continuously developing constructs of primary interest in the education sciences and outline key advantages of using this type of model. Furthermore, we propose the *continuous-time latent curve model with structured residuals* (CT-LCM-SR) as a suitable model for many research questions in the education sciences. The CT-LCM-SR combines growth and dynamic modeling and thus provides descriptions of both trends and process dynamics. We illustrate the application of the CT-LCM-SR with data from PISA reading literacy assessment of 2000 to 2018 and provide a tutorial and annotated code for setting up the CT-LCM-SR model.

**Keywords:** Longitudinal data analysis, Continuous-time, Dynamic model, Latent curve models, PISA

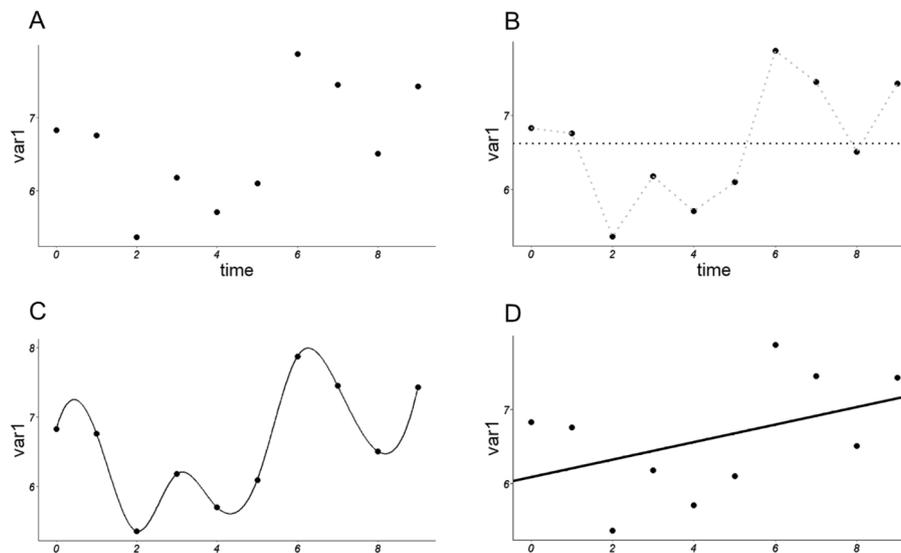
## Introduction

Longitudinal studies and methods are essential to many research questions in education sciences. Questions of learning, growth, and development are studied by examining change over time (Millsap, 2008; Singer, 2019). Moreover, if researchers are interested in causal relationships between several constructs, the temporal order, as realized in longitudinal studies, is one important prerequisite (among others) for drawing robust conclusions (Shadish et al., 2002). One challenge associated with longitudinal studies is the choice of the appropriate statistical model for the data analysis. The statistical models should correspond with the research questions at hand and the theoretical assumptions about the constructs and the processes underlying them (Baltes & Nesselrode, 1979; Collins, 2006; Little, 2013; McArdle, 2009). A close fit between theories of change and the models used to analyze the data is a prerequisite for valid inferences.

One class of models that is frequently applied in educational research is dynamic models, which include, for example, (vector) autoregressive models (e.g., Hsiao, 2014; Lütkepohl, 2005), cross-lagged panel models (e.g., Hamaker et al., 2015; Usami et al., 2019; Zyphur, Allison, et al., 2020; Zyphur, Voelkle, et al., 2020), or latent change score models (McArdle & Hamagami, 2001). These models usually treat time as discrete, which is why they are often called discrete-time dynamic models (DTMs; e.g., Hsiao, 2014; Voelkle et al., 2018). DTMs, however, are limited by the widely recognized *time-interval dependency* (e.g., Gollob & Reichardt, 1987; Ryan & Hamaker, 2021). This means that dynamic parameters of DTMs depend on a respective time-interval. Thus, they cannot represent the continuous nature of developmental processes and do not provide information regarding other time intervals, which is a considerable limitation in several other regards (e.g., Voelkle et al., 2012). For this reason, continuous-time dynamic models (CTMs) have been put forward in recent years (van Montfort et al., 2018).

In the present study, we adopt these ideas and outline potential advantages of CTMs over DTMs for longitudinal data analysis in educational research. In addition, we focus on how trends can be integrated into dynamic models. Besides the analyses of process dynamics, trends are additional characteristics of longitudinal data and are often of core interest in educational research. Several models have been proposed to capture trends and dynamics simultaneously and are discussed in ongoing debates (e.g., Asparouhov et al., 2018; Hamaker, 2005; Núñez-Regueiro et al., 2021; Usami et al., 2019). However, to date, most of them employ discrete-time modeling to represent the dynamic part (but see, e.g., Delsing & Oud, 2008). Dynamic models that provide information on trends and dynamics on a continuous-time scale, would overcome this limitation. In the present article, we propose a combined model that provides information on both the continuous-time dynamics and trends that we call the *continuous-time latent curve model with structured residuals* (CT-LCM-SR). The model is a continuous time version of the existing *latent curve model with structured residuals* (LCM-SR; Curran et al., 2014; Hamaker, 2005). We illustrate the application of the CT-LCM-SR with data from the PISA reading literacy assessment using the R package *ctsem* (Driver & Voelkle, 2018a, 2021). In doing so, two research questions are examined. (1) Is there a trend in the PISA countries' mean reading scores over the period from 2000 to 2018, and do the countries differ in the slope and direction of the trend? (2) How persistent are deviations from the trend? The latter research question tackles the stability of educational systems when deviations from the trend (so-called shocks or impulses) occur. Influences leading to temporary gains (positive deviations from the trend) could be, for example, temporary investments or short-term interventions initiated by a country's politics. Influences which might result in negative deviations can be political unrest, a shortage of teachers in a given cohort, or a global financial crisis or pandemic. Thus, the stability of a system could be thought of as the system's resilience against disturbances. Clearly, this is oftentimes a double-edged sword, as a high persistence of positive deviations is often desired, which would be the case in an instable or non-resilient system. Contrarily, negative disturbances are often desired to dissipate quickly, which would be the case if the system is stable or resilient.

In our illustration, we show how the CT-LCM-SR can be used to answer both research questions. Thereby, we compare the CT-LCM-SR to related models, exemplarily



**Fig. 1** Differences between measurement of change (**A**), modeling of change (**B, D**) and theories of change (**C**). Panel **A** Simulated data for a single observation unit examined at ten measurement occasions. Panel **B** DT autoregressive effects are usually interpreted by reference to a constant process mean. The process mean is represented by the horizontal dotted line. Panel **C** In theories of change, it is usually assumed that there is continuous existence of observation units and continuous change of variables, which is represented by the solid line connecting successive measurements. Panel **D** A linear trajectory smoothed through the data, as is done in (linear) growth-curve models. The model-implied trajectory interpolates any time point and can be used to predict values for any time-interval using time as a continuous predictor

interpret the parameter estimates of the CT-LCM-SR, and present some details and ways of visualizing results to emphasize the advantages of CTMs over DTMs.

This article is aimed at two audiences. In the first section, we introduce CTMs in general and give arguments for why the use of CTMs can be beneficial. This section may be particularly interesting to readers who are new to continuous-time modeling. For readers who are already familiar with this type of modeling, we provide detailed information on the implementation and interpretation of the newly proposed CT-LCM-SR model in the methods section. Furthermore, readers get a practical introduction to the use, modeling, and interpretation of CTMs. Finally, the strengths and limitations of the illustrative example in particular and CTMs in general are discussed. The discussion also includes an outlook on more complex (multivariate) applications of the CT-LCM-SR.

### How variables develop over time and how we measure this development

When longitudinal studies are conducted in educational research, the data are usually repeated measurements of the same constructs and target units (e.g., subjects). For instance, questionnaires or achievement tests are repeated multiple times, each after a certain period of time. Based on such snap-shots in time, we gain insights about the state of our target units (Fig. 1A). Between these measurement occasions, we have no data and thus no information about the changes of the variables under study. The stability, inertia, development, change, or longitudinal association of multiple variables in multivariate analysis must be inferred from these repeated snapshot-like data. In many theories of change, however, the observation units and variables of interest are not assumed to exist only at the measurement occasions or to change in discrete steps (Fig. 1B). Rather,

we assume that observation units and variables under investigation exist and develop continuously over time (Fig. 1C). This is applicable to variables at the level of individuals, such as the development of students' competencies, cognitive abilities, self-concepts, attitudes, and motivations. Moreover, this continuous-time perspective is transferable to social systems and variables at the level of school classes, variables describing parent-child-teacher interactions, or the performance of entire education systems as studied, for instance, in the international large-scale assessments (LSAs) such as PISA and TIMSS. Although there are typically several years between each measurement occasion in LSAs, we do not assume that the characteristics of education systems cease to exist in the interim. Thus, from a theoretical point of view, it is desirable to adopt the idea of continuously developing variables also in the statistical models used. As shall become apparent in the following sections, this not only allows us to align models closer with theories of change, but also provides further advantages, such as greater flexibility for the procedures of data collection, improvements for cross-study comparisons (e.g., Voelkle et al., 2012), or the exploration of the unfolding of dynamic effects (e.g., Hecht & Zitzmann, 2021a).

#### Discrete-time dynamic models (DTMs)

Dynamic models represent the state of variables at time  $t_i$  as a function of previous states (e.g., at time  $t_{i-1}$ ) of the same variables (Deboeck & Boker, 2015; Hsiao, 2014; Voelkle et al., 2018). Typical questions that can be addressed with dynamic models are questions related to the stability or inertia of constructs over time, to find out, for example, how fast the states of variables change (or can be changed). Dynamic models can also be used to address questions about the variability of states of a construct or questions about the predictability of future states (based on the current state) (Deboeck & Boker, 2015; Loosens et al., 2021). Well-known dynamic models in educational research are multivariate models, such as the cross-lagged panel models (CLPM). In terms of *Granger causality*, such models usually aim to provide insights into the strength and predominant direction of the relationship between the variables of interest (Granger, 1969; Hamaker et al., 2015).

In educational research, many dynamic processes are modeled by first-order models such as the first-order autoregressive model (AR(1)) or its multivariate variant, the vector autoregressive model (VAR(1))<sup>1</sup> (Sivo & Fan, 2008). In VAR(1) models, the present state of variables is predicted by the previous state. A discrete-time VAR(1) model can be represented by the following equation (cf. Oud & Delsing, 2010; Voelkle et al., 2012):

$$\mathbf{x}(t_i) = \mathbf{A}_{\Delta t_i} \mathbf{x}_{t_i - \Delta t_i} + \mathbf{b}_{\Delta t_i} + \mathbf{w}_{t_i - \Delta t_i} \text{ with } (\mathbf{w}_{t_i - \Delta t_i}) \sim N(0, \mathbf{Q}_{\Delta t_i}). \quad (1)$$

Here,  $\mathbf{x}$  is a column vector of length  $V$  (number of variables), observed at time  $t_i$  (with  $i = 1, \dots, T$ , where  $T$  is the number of measurement occasions) and  $\Delta t_i$  the time interval between two measurement occasions.  $\mathbf{A}$  is a  $V \times V$  matrix relating the observed variables over time (autoregressions on the main diagonal, cross-lagged effects on the

<sup>1</sup> The term AR or VAR model stems from the tradition of time series analysis. In classical time series analyses, only a single unit ( $N=1$ ) is often examined (e.g., Ryan et al., 2018). The term CLPM can be understood as a label for a version of the VAR model for panel data (cf. Kuiper & Ryan, 2018), with the caveat that sometimes, the lagged coefficients are not equated across measurement occasions in CLPMs (Sivo & Fan, 2008). Because panel data are the norm in educational research, the term VAR model is used in the following to refer to the *panel VAR model* (Hsiao, 2014; Sivo & Fan, 2008).

off-diagonals),  $\mathbf{b}$  is an intercept vector of length  $V$ , and  $\mathbf{w}$  a vector of prediction errors or white noise (e.g., Deboeck & Preacher, 2016; Oud & Delsing, 2010). Covariance matrix  $\mathbf{Q}_{\Delta t_i}$  contains prediction error variances and covariances. From this representation of a VAR(1) model, it is apparent that the coefficients of interest, in particular the autoregressive and cross-lagged coefficients of the matrix  $\mathbf{A}$ , depend on the time-interval  $\Delta t_i$  ( $\Delta t_i = t_i - t_{i-1}$ ) between successive measurement occasions. For example, if this interval length is three months, the estimated coefficients are conditional on exactly this time-interval and are not valid for any other interval. Thus, the process described by the parameters corresponds to the development of a variable or an observation unit that exists only at the time points of measurement, illustrated in Fig. 1A. However, as discussed earlier, in most cases this representation is not consistent with theory that assumes that variables and observation units exist and change continuously over time (Fig. 1C). This *time-interval dependency* is also referred to in literature as the *interval problem* of DTMs (e.g., Ryan & Hamaker, 2021). This dependency results in several drawbacks, which have been described in detail elsewhere (e.g., Hecht et al., 2019; Oud & Delsing, 2010; Voelkle et al., 2012) and will only be briefly mentioned here: (1) The information about the dynamics of the variables under study that DTMs reveal only relates to the particular time-interval used (e.g., Deboeck & Preacher, 2016; Kuiper & Ryan, 2018). (2) There is no information about how the dynamic effects change with increasing time-intervals (e.g., Hecht & Zitzmann, 2021a). Moreover, (3) (approximately) equal spacing between measurement occasions must be realized across all units in order to be able to properly apply DTMs<sup>2</sup> (see Driver et al., 2017). Establishing equal spacing is a major challenge, especially when using flexible survey designs (de Haan-Rietdijk et al., 2017) whose importance has been increasingly emphasized for educational research in recent years (Zirkel et al., 2015). The time-interval dependency of DTMs also has disadvantages since (4) results from studies examining different time-intervals between measurement occasions are not directly comparable with each other, making it difficult to conduct meta-analyses on these results (Dormann et al., 2020). Continuous-time dynamic models (CTMs) resolve these problems.

### Continuous-time dynamic models (CTMs)

The assumption of variables that exist and unfold continuously is consistent with CTMs (e.g., Kuiper & Ryan, 2018; Oud & Delsing, 2010). By using CTMs, we obtain not only the same information provided by DTMs if all underlying assumptions are met, but also parameters that are independent of the length of the time-intervals of a given study. These continuous-time parameters allow for gaining insights into the underlying continuous-time process and for deriving discrete-time parameters for various time-intervals (Hecht et al., 2019; Voelkle et al., 2012). In addition, CTMs overcome the practical drawbacks and limitations of DTMs outlined above (e.g., Oud & Delsing, 2010; Voelkle & Oud, 2013). However, the implementation of CTMs is mathematically more challenging than that of DTMs because CTMs require differential calculus (Deboeck & Boulton, 2016; Voelkle et al., 2018). Fortunately, the application of CTMs has been greatly facilitated in recent years by software

<sup>2</sup> One possible solution to address unequally spaced measurement occasions using DTMs is the so-called *phantom-variable approach*, but this approach is limited to few unequal intervals. Furthermore, other drawbacks associated with the interval-dependency of DTMs remain (see Oud & Voelkle, 2014).

packages like the R package *ctsem* (Driver et al., 2017; Driver & Voelkle, 2017, 2018a, 2021). As a result, the application of a wide range of CTMs is now possible in a relatively straightforward manner.

The basic stochastic differential equation (SDE) used in *ctsem* can be represented as follows (Oud & Delsing, 2010; Voelkle et al., 2012; for an overview of different notation styles of the same equation see, e.g., Hecht & Zitzmann, 2021a):

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{A}\mathbf{x}(t) + \mathbf{b} + \mathbf{G}\frac{d\mathbf{W}(t)}{dt}. \quad (2)$$

The derivative  $\frac{d\mathbf{x}(t)}{dt}$  provides information about the rate of change of the variables in the column vector  $\mathbf{x}$  of length  $V$  at the time  $t$  (Deboeck & Boulton, 2016; Ryan et al., 2018). The term can be thought of as the change  $dx$  in the variable  $x$  over the time-interval  $dt$ , which, in case of the derivative, is infinitesimally small ( $\Delta t_i \rightarrow 0$ ). The expression  $\frac{d\mathbf{x}(t)}{dt}$  is also referred to as *velocity* (Deboeck et al., 2015). Given the information regarding the rate of change at time  $t$ , we know how the process under study is changing at that time (Ryan et al., 2018). The rates of change of the variables  $\mathbf{x}$ , shown on the left side of Eq. 2, is explained by a deterministic and a stochastic part on the right side. The deterministic part is composed of the so-called drift matrix  $\mathbf{A}$ , which is a  $V \times V$  matrix and includes the continuous-time auto-effects (equivalent to autoregressive effects in DTMs but independent from the length of time-intervals) on the main diagonal and the continuous-time cross-effects on the off-diagonals, the state of the variables  $\mathbf{x}$  at time  $t$ , and a  $V \times 1$  continuous-time intercept vector  $\mathbf{b}$ .  $\mathbf{G}\frac{d\mathbf{W}(t)}{dt}$  represents the stochastic part of the equation and can be described as a continuous-time error process (Oud & Delsing, 2010; Ryan et al., 2018).  $\mathbf{W}(t)$  is the so-called Wiener process, which is a random walk in continuous-time and the Cholesky factor  $\mathbf{G}$  indicates the effect of  $\mathbf{W}$  on the change in  $\mathbf{x}(t)$  (for details see Driver & Voelkle, 2018a; Oud & Jansen, 2000). Because the variance of the Wiener process depends on the length of the time-interval over which it is integrated, it can represent the accumulation of errors over longer time periods, resulting in larger error variances (Deboeck & Preacher, 2016; Oravecz et al., 2011). The associated  $V \times V$  variance–covariance matrix  $\mathbf{Q} = \mathbf{G}\mathbf{G}'$  is often called *diffusion matrix* and contains the process error variances on the main diagonal and the process error covariances on the off-diagonals.

Based on the continuous-time (CT) parameters of Eq. 2 it is possible to calculate the DT parameters for any given time-interval (Hecht et al., 2019; Oud & Delsing, 2010). To better distinguish between CT parameters and its corresponding DT parameters, in the following, DT parameters that are constrained to underlying CT parameters are denoted with an asterisk (\*) (cf. Hecht et al., 2019). For an overview of corresponding DT and CT parameter labels, see also Hecht and Voelkle (2021). The CT drift matrix  $\mathbf{A}$  is related to the DT autoregressive matrix  $\mathbf{A}_{\Delta t_i}^*$  via the following equation:

$$\mathbf{A}_{\Delta t_i}^* = e^{\mathbf{A}\Delta t_i}. \quad (3)$$

The CT intercept vector  $\mathbf{b}$  is related to the DT intercept vector  $\mathbf{b}_{\Delta t_i}^*$  via the following equation:

$$\mathbf{b}_{\Delta t_i}^* = \mathbf{A}^{-1} \left( e^{\mathbf{A}\Delta t_i} - \mathbf{I} \right) \mathbf{b}. \quad (4)$$

Furthermore, the DT error covariance matrix  $\mathbf{Q}_{\Delta t_i}^*$  for interval  $\Delta t$  can be calculated from the CT diffusion matrix  $\mathbf{Q}$  (Delsing & Oud, 2008):

$$\mathbf{Q}_{\Delta t_i}^* = \text{row} \left[ (\mathbf{A} \otimes \mathbf{I} + \mathbf{I} \otimes \mathbf{A})^{-1} (\mathbf{A}_{\Delta t_i}^* \otimes \mathbf{A}_{\Delta t_i}^* - \mathbf{I} \otimes \mathbf{I}) \text{row}(\mathbf{Q}) \right], \quad (5)$$

where  $\mathbf{I}$  is an identity matrix,  $\otimes$  the Kronecker product, which results in a large matrix containing all possible products of the elements of the two initial matrices, *row* and *row* are operators, the former puts the elements row-wise in a column vector and the latter puts them from a column vector into a matrix. To solve the system of Eqs. 2, 3, 4, 5, the matrix exponential has to be used (e.g., Oud et al., 2018; Ryan et al., 2018). Details and examples of the link between the parameters and their interpretation will be described in the methodical and results section of the article. Before doing so, however, we address another challenge that often arises when analyzing longitudinal data in education sciences: systematic trends.

### Trends

Analyzing longitudinal data in educational research is often also about identifying trends. For example, when studying students' achievement repeatedly over several months and years, their achievement typically continues to increase. Such systematic trends can also occur when examining other variables, such as motivational (e.g., Wigfield et al., 2006, 2015) or personality-related characteristics (e.g., Specht, 2017). Furthermore, besides the theoretical interest in trends, it is also often important for dynamic models to account for such trends. Unaccounted trends can affect the autocovariance structure and thus lead to biased estimates of the parameters (e.g., Asparouhov et al., 2018; Núñez-Regueiro et al., 2021; Walls & Schäfer, 2006). Therefore, there are several approaches to deal with trends in dynamic models, such as pre-detrending the data (e.g., Box et al., 2015; Walls & Schäfer, 2006). Another possibility is to model trends and dynamics simultaneously. *Growth-curve models* (GCM) are a well-known approach to model trends (Fig. 1D) (e.g., Bollen & Curran, 2006). There are approaches that combine dynamic DTMs with GCMs, such as the *autoregressive latent trajectory* (ALT) models (e.g., Bianconcini & Bollen, 2018; Bollen & Curran, 2006; Curran et al., 2014; Hamaker, 2005). In addition to all the other disadvantages of DTMs already described, approaches combining elements from DTMs and GCMs come with a model-inherent unequal time-interval dependency of parameters: While GCM-parameters are usually treated and interpreted as independent from time-intervals (Fig. 1D), the DTM parameters depend on the time-interval used in the study (Delsing & Oud, 2008). Thus, in combined models such as ALT models, the trend parameters describe a continuous process (Fig. 1D) and the dynamic parameters describe a discrete-time process (Fig. 1B). To remove this inconsistency from combined models, a promising solution is to use CTMs to make the dynamic parameters independent of time-intervals as well. One of the very first explicit formulations of a combined model in continuous-time is the *continuous-time autoregressive latent trajectory* (CALT) model proposed by Delsing and Oud (2008; see also Oud, 2010).

However, ALT and CALT models are associated with other considerable limitations. Trends and dynamics are not clearly separated in ALT and CALT models and can

influence and “compete” with each other. An associated problem is that dynamic models are recursive, which means that dynamic processes depend on their previous values. This is not the case for GCMs. These problems lead to drawbacks, such as the need for special solutions to “start up” ALT and CALT models and uncertainty of how to interpret GCM parameters in ALT and CALT models (e.g., Bianconcini & Bollen, 2018; Hamaker, 2005; Jongerling & Hamaker, 2011; Little, 2013; Oud, 2010; Usami et al., 2019).

To clearly isolate trends from the dynamic part of the model, the discrete-time (DT) *latent-curve model with structured residuals* (LCM-SR) has been suggested as one suitable alternative (Curran et al., 2014; Hamaker, 2005). In the LCM-SR, the development of a variable under study is described by two different process components, trends and dynamics, which are—contrarily to the ALT or CALT models—clearly separated (Fig. 2). Because the trend component (which is similar to what is modeled in the GCM) accounts for the trends in the data, the dynamic component represents a trend-adjusted (detrended) and centered process in the LCM-SR (Usami et al., 2019). For this reason, systematic (linear) trends are disentangled from the autoregressive and cross-lagged parameters. An additional advantage of this separation is that the parameters of both processes can be interpreted as is typically done in CTMs and GCMs. Moreover, trend parameters and dynamic parameters do not compete in a joint process in the LCM-SR as in the ALT and CALT models (see Curran et al., 2014; Hamaker, 2005; Jongerling & Hamaker, 2011).<sup>3</sup>

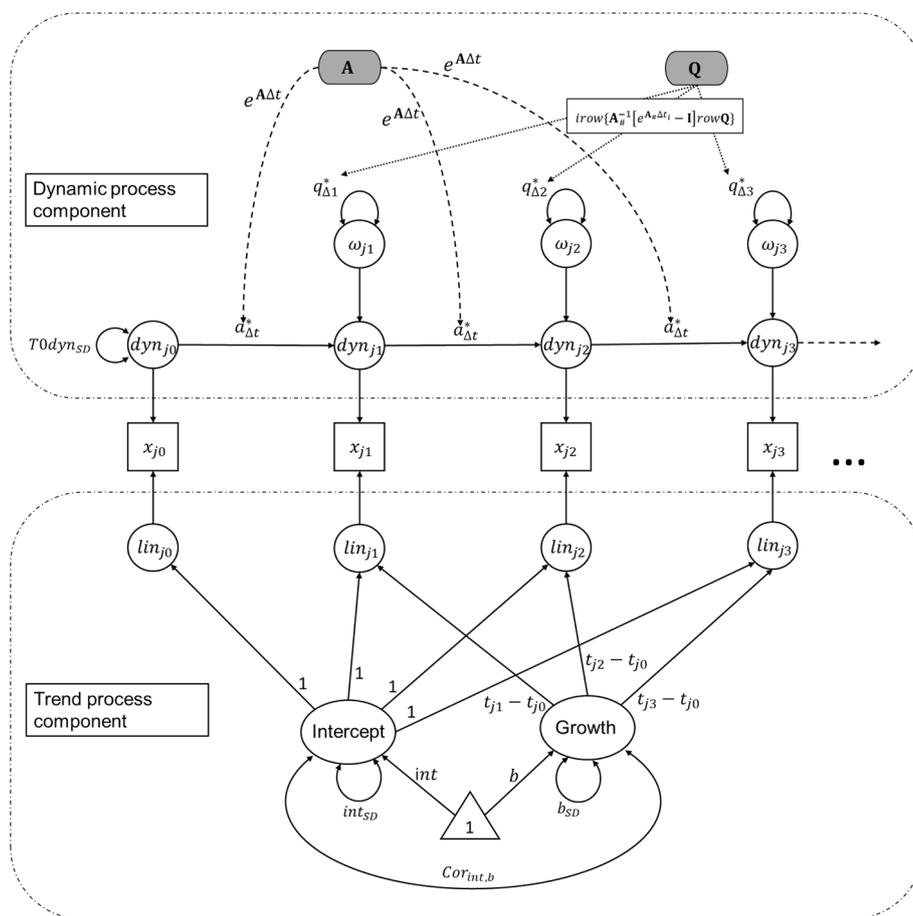
To our knowledge, an explicit formulation of a continuous-time version of the LCM-SR has not been proposed in the literature, although a *continuous-time latent curve model with structured residuals* (CT-LCM-SR; Fig. 2) comes with several advantages: (1) Using the CT-LCM-SR, information is obtained describing trends in the data as well as information describing the process’ dynamics on a continuous-time scale. (2) Trends and dynamic processes are clearly separated in CT-LCM-SR, meaning that trends and dynamics can no longer influence each other. (3) The interpretation of the parameters follows the interpretation of CTMs and GCMs. (4) All parameters of the CT-LCM-SR describe the process under study independently of the length of time-intervals in a given study. This is not the case for the DT-LCM-SR in which the dynamic parameters always depend on the length of the time-interval (cf. Delsing & Oud, 2008; Oud, 2010). Finally, (5) both the standard linear GCM and the CT-AR(1) model (which is a CTM without trend components) are nested within the CT-LCM-SR which enables chi-square difference testing to compare models. In addition, with the CT version of the LCM-SR, we can take advantage of all the other benefits associated with using CTMs instead of DTMs.

## Methods

### The CT-LCM-SR

The basic idea of the CT-LCM-SR is to decompose a process into the two process components, trends and dynamics. This is technically done by introducing two continuous-time processes with special parameter values and constraints, so that one continuous-time process purely captures the trend and the other one the dynamics (see technical details below, especially Eq. 8 and 9). In the following, we therefore also refer

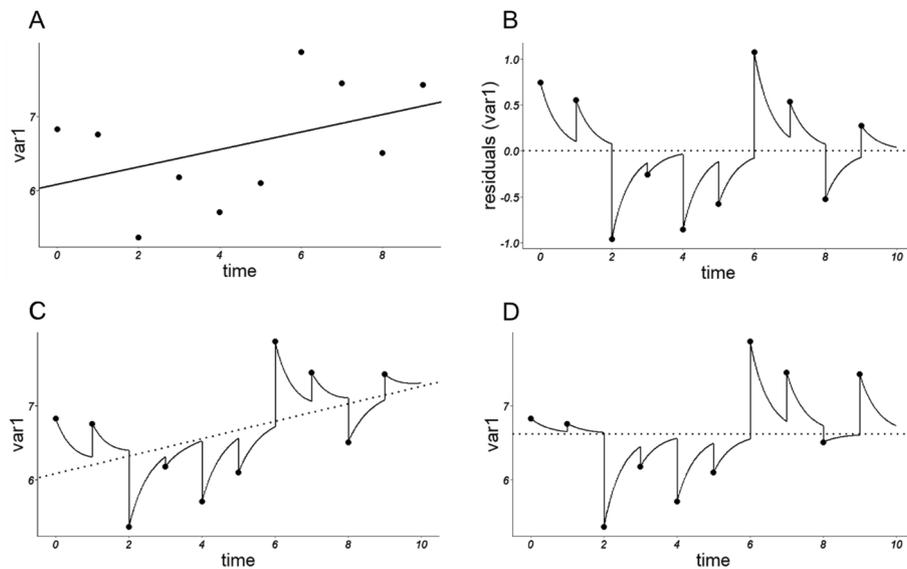
<sup>3</sup> Curran et al. (2014) also elaborate in detail on the advantage of disentangling between-unit differences in trends from the dynamics taking place at the within-unit level (see also Núñez-Regueiro et al., 2021).



**Fig. 2** Continuous-time Latent Curve Model with Structured Residuals. The CT-LCM-SR with four measurement occasions: Manifest scores  $x_{jt}$  are represented by the squares. On the bottom of the squares, there is the trend process (standard linear growth-curve process with random effects for intercept and slope) accounting for differences in the initial values and trends. The continuous-time dynamic process, which accounts for the dynamic structure is represented on the top of the squares. Estimated parameters:  $int$  mean intercept,  $int_{SD}$  variance of the random effects of the intercept,  $b$  mean growth rate,  $b_{SD}$  SD of the random effects of the growth rate,  $Cor_{int,b}$  intercept-growth correlation,  $T0dyn_{SD}$  initial SD of the dynamic process;  $A$  drift matrix,  $Q$  diffusion matrix

to the trend component as the “trend (or growth) process” and to the dynamic component as the dynamic process with the need to keep in mind that these are just technical terms for the two components of one single process variable.

The univariate CT-LCM-SR is thus composed of two process components, one to describe the linear trend and one to model the dynamics of one process variable. When including more variables, each variable would need two processes to represent its two process components, for instance, a bivariate CT-LCM-SR would then include four processes, a trivariate CT-LCM-SR six, and so forth. The ordering of the two processes is arbitrary. We use the first process as the linear *growth process* that accounts for initial values and trends (Fig. 3A). The second process is then the *dynamic process* that provides



**Fig. 3** Combination of trend (A) and dynamics (B) in the CT-LCM-SR (C) in comparison to a CT-AR(1) model (D). Same simulated data as used in Fig. 1. Panel A CT-LCM-SR-implied trend. Panel B continuous-time dynamic process of the residuals. Panel C Putting trend and dynamic process together. Panel D A standard CT-AR(1) model, which ignores any trend in the data. Black solid lines visualize model implied expected trajectories (given past observations). The model implied predictions of the CT-LCM-SR (Panel C) are a combination of the static trend component (Panel A) and the autoregressive structure of the residuals over time as reflected by the CT auto-effect (Panel B). The dynamic part gives information on how much of the deviation from the trend on time  $T$  is expected to be carried over to future states with growing time intervals. A mean-reverting process is shown, which means that the CT auto effect is negative (corresponding to DT autoregressive effects between 0 and 1 for every time interval)

information about the process' dynamics on a continuous-time scale (Fig. 3B).<sup>4</sup> Since the growth process accounts for nonzero initial values (intercept parameter) and for linear trends (growth parameter), the dynamic process can be considered as centered and linearly detrended (cf. Usami et al., 2019). By combining the two process components in one model, the values can be thought of as fluctuating around the linear trend (Fig. 3C).

Just as in growth-curve modeling, CT-LCM-SR can account for differences between observation units in intercept and growth parameters via random effects (e.g., Bollen & Curran, 2006; Little, 2013). Thus, the dynamic process is within-unit centered. If potential differences in trends and initial values between observation units are not taken into account, within and between level effects can be confounded (see also the discussion of the use of random effects in dynamic models by, e.g., Hamaker et al., 2015; Lüdtke & Robitzsch, 2021; Núñez-Regueiro et al., 2021; Usami et al., 2019).

Using the SDE (Eq. 2), growth-curve models can be specified as special cases (Driver, 2020). To illustrate the application of a CT-LCM-SR with the SDE, we first consider a single growth-curve process.

To obtain a GCM with the SDE, the CT auto-effect is set to a negative value very close to zero (e.g.,  $-0.0001$ ). This means, that the corresponding DT autoregressive effect ( $a_{\Delta t_i}^* = e^{a\Delta t}$ ) approximates 1. The freely estimated initial value  $x_0$

<sup>4</sup> Figure 3B shows a so called mean-reverting process, which means, that the CT auto-effect is negative (i.e., discrete-time autoregressive coefficients are between 0 and 1 for all time intervals). Other types of dynamics are also possible, but they are rare in educational and psychological research (cf. Hamaker et al., 2015), which is why we are limiting ourselves to this type of dynamics in the present article.

corresponds to the growth-curve intercept parameter, while the continuous-time intercept ( $b_{\Delta t_i}^* = a^{-1}(e^{a\Delta t} - 1)b$ ) serves as the linear growth parameter. To illustrate the functionality, we consider the deterministic part of Eq. 2 and insert the DT parameter constraints of Eqs. 3 and 4 (Oud & Delsing, 2010):

$$E[x_t] = e^{a\Delta t}E[x_0] + a^{-1}(e^{a\Delta t} - 1)b. \tag{6}$$

The right part of the equation is composed of the two constraints relating the CT auto-effect to the DT autoregressive effect and the continuous-time intercept to the DT intercept.  $E[x_0]$  is the expected value of the freely estimated initial mean. To obtain the expected values for other measurement occasions, the value for  $\Delta t$  is changed. To illustrate this, let us assume  $\Delta t = t_1 - t_0 = 1$  for the second measurement occasion  $t_1$ :

$$\begin{aligned} E[x_1] &= e^{-.0001 \times 1}E[x_0] + -.0001^{-1}(e^{-.0001 \times 1} - 1)b \\ E[x_1] &= 1 \times E[x_0] + 1 \times b \end{aligned} \tag{7}$$

Compared to the initial measurement occasion, the expected value for the second measurement occasion increases by the additive component  $1 \times b$ . To obtain the expected value for the third measurement occasion  $t_2$ , we set  $\Delta t = 2$ :

$$E[x_2] = e^{-.0001 \times 2}E[x_0] + -.0001^{-1}(e^{-.0001 \times 2} - 1)b$$

$$E[x_2] = 1 \times E[x_0] + 2 \times b$$

The expected value for  $t_2$  increases by the additive component  $2 \times b$  compared to the initial measurement occasion. This can be continued for any measurement occasion. Because the CT auto-effect is fixed to a negative value close to zero, which corresponds to a DT autoregressive effect approximating 1, the continuous-time intercept part linearly increases with longer time-intervals.

To obtain a univariate CT-LCM-SR (Fig. 2), we extend Eq. 6 by a second (dynamic) process with a freely estimated auto-effect  $a$  and a process mean of zero ( $b = 0$ ):

$$\begin{aligned} E \begin{bmatrix} x_{lint} \\ x_{dyn_t} \end{bmatrix} &= \begin{bmatrix} e^{-.0001\Delta t} & 0 \\ 0 & e^{a\Delta t} \end{bmatrix} \times \begin{bmatrix} x_{lint-\Delta t} \\ x_{dyn_{t-\Delta t}} \end{bmatrix} (t) \\ &+ \begin{bmatrix} -.0001^{-1}(e^{-.0001\Delta t} - 1) & 0 \\ 0 & a^{-1}(e^{a\Delta t} - 1) \end{bmatrix} \times \begin{bmatrix} b \\ 0 \end{bmatrix} \end{aligned} \tag{8}$$

$$\text{with } x_t = [\lambda_1 \lambda_2] \begin{bmatrix} x_{lint} \\ x_{dyn_t} \end{bmatrix} = [11] \begin{bmatrix} x_{lint} \\ x_{dyn_t} \end{bmatrix} = x_{lint} + x_{dyn_t}. \tag{9}$$

The lambda matrix  $\lambda = [\lambda_1 \lambda_2]$  in Eq. 9 relates the observed values of variable  $x$  at time  $t$  to the two processes  $x_{lint}$  and  $x_{dyn_t}$ , which represent the growth component and the dynamic component, respectively. The states of the two processes  $x_{lin}$  and  $x_{dyn}$  at time  $t$  are predicted by the state of the two processes at a previous time point  $t - \Delta t$  (Eq. 8). The autoregressive matrix is the first element on the right side of the equation. On the main diagonal, the autoregressive matrix has the restricted parameter  $e^{-.0001\Delta t}$  for the linear trend component and the freely estimated autoregression  $e^{a\Delta t}$  for the dynamic

component. The off-diagonals are set to zero since the two process components are not assumed to directly affect each other. The intercept vector (the second element on the right-hand side of Eq. 8) contains the known restrictions for the growth process component at the top, and is set to zero at the bottom, because the dynamic process component is centered and detrended (fluctuating around the process-mean of zero).

The model can be represented in the SDE (Eq. 2) notation as follows:

$$d \begin{bmatrix} x_{lin} \\ x_{dyn} \end{bmatrix} (t) = \left( \begin{bmatrix} -.0001 & 0 \\ 0 & a \end{bmatrix} \times \begin{bmatrix} x_{lin} \\ x_{dyn} \end{bmatrix} (t) + \begin{bmatrix} b \\ 0 \end{bmatrix} \right) dt + \begin{bmatrix} 0 & 0 \\ 0 & g \end{bmatrix} \times \begin{bmatrix} 0 \\ w_{dyn} \end{bmatrix}. \quad (10)$$

Here, it becomes obvious that the trend component is deterministic because the random error term is zero. Therefore, the residuals are completely absorbed into the dynamic part and are accounted for by CT auto-effect  $a$  and CT random error variance  $Q$ . In addition, the first measurement occasion ( $T0$ ) is estimated as

$$\begin{bmatrix} x_{lin_{t0}} \\ x_{dyn_{t0}} \end{bmatrix} \sim N \left( \begin{bmatrix} int_{MEAN} \\ 0 \end{bmatrix}, \begin{bmatrix} int_{SD} & 0 \\ 0 & T0_{dyn_{SD}} \end{bmatrix} \right), \quad (11)$$

where  $int_{MEAN}$  represent the mean intercept of the trend component,  $int_{SD}$  the standard deviation ( $SD$ ) of the random effects of the intercept, and  $T0_{dyn_{SD}}$  the residual  $SD$  of the first measurement occasion entering the dynamic process (the labels of the estimated parameters in Eq. 10 and 11 correspond to the labels used in the illustrated example later on and Fig. 2). In the random effects version of the CT-LCM-SR, there is an additional variance component for the slope of the trend component  $b \sim N[b_{Mean}, b_{SD}]$  and a correlation of the random effects ( $Cor_{int,b}$ ).

To turn this CT-LCM-SR into a model without trend component (i.e., a standard CT-AR(1) model), we can set the freely estimated parameter  $b$  in Eq. 10 to zero.

### An illustrative example of the CT-LCM-SR model using PISA data

In the following, we provide an illustrative application of the CT-LCM-SR to data from the PISA reading literacy assessment, including a step-by-step tutorial on how to set up the CT-LCM-SR in *ctsem*. To make it as understandable as possible, a simple example was chosen, which can, however, easily be extended to more complex models and research questions (as outlined in the outlook of this article).

### Data

The data stem from the seven currently available PISA waves from 2000 to 2018. For each single wave, country-specific mean scores were calculated across all participating students in each country (see OECD, 2009). For this step of data preparation, the R-package *intsvy* (Caro & Biecek, 2017) was used and the resulting values were additionally checked with the SPSS macros provided on the OECD website. This was primarily done to ensure that the weighting procedures were accurate (IBM SPSS, 2020; OECD, 2021). Furthermore, because data from the achievement tests were processed with the Rasch model, the use of *plausible values* (PVs) is recommended (OECD, 2009;

	<i>id</i>	<i>time</i>	<i>wave</i>	READ	CNT
1	1	0	2000	348.43	ALB
2	1	3	2003	NA	ALB
3	1	6	2006	NA	ALB
4	1	9	2009	385.21	ALB
5	1	12	2012	393.97	ALB
6	1	15	2015	404.97	ALB
7	1	18	2018	405.84	ALB
8	2	0	2000	414.64	ARG
9	2	3	2003	NA	ARG
10	2	6	2006	373.71	ARG
11	2	9	2009	398.63	ARG
12	2	12	2012	395.39	ARG
13	2	15	2015	NA	ARG
14	2	18	2018	402.21	ARG
15	3	0	2000	529.15	AUS
16	3	3	2003	525.04	AUS
17	3	6	2006	513.11	AUS
18	3	9	2009	514.75	AUS
19	3	12	2012	511.47	AUS
20	3	15	2015	502.40	AUS
21	3	18	2018	503.42	AUS
22	4	0	2000	509.15	AUT
23	4	3	2003	490.50	AUT
24	4	6	2006	490.49	AUT
25	4	9	2009	471.02	AUT
26	4	12	2012	489.96	AUT
27	4	15	2015	484.91	AUT
28	4	18	2018	484.55	AUT
29	5	0	2000	509.02	BEL
30	5	3	2003	507.33	BEL

**Fig. 4** PISA data structured in long format. The column *id* assigns the measurements to the countries creating the hierarchical data structure. The information of when a respective measurement took place is contained in the *time* column

Wu, 2005). Therefore, all analyses were first performed with each PV separately and the results were pooled afterwards (OECD, 2009; Rubin, 2004). In PISA, the achievement data are represented by five PVs.<sup>5</sup> Thus, five data sets were created, one for each PV (see OECD, 2009).

The data sets had a hierarchical data structure, with repeated measurements nested within countries. For analyses with the *ctsem* package (Driver & Voelkle, 2017), the data sets should preferably be in the *long format* (see also the *ctsem* tutorial by Hecht et al., 2019). Therefore, the data were structured, with all observations of a variable being contained in the same column (in contrast to the *wide format*, where there is a separate column for each measurement occasion). The corresponding time of the measurement was put in a separate column. Information about which observations were related to which country was contained in another column (see Fig. 4).

### Sample

The final sample examined in this study consisted of  $N = 56$  PISA countries. It is the same sample selected by the OECD (2019a) for the presentation of countries developmental

<sup>5</sup> Since PISA 2015, there is a shift to including ten PVs. Due to the smaller number of five corresponding PVs in the previous waves, only the first five PVs were used in the data preparation for the 2015 and 2018 waves in the present study.

trajectories in mean reading literacy performance. To be included, a country had to have participated in at least five PISA surveys between 2000 and 2018.

### **Measures**

*PISA reading literacy* The primary variable of interest are the countries' mean estimates in PISA reading literacy assessment. Because the reading scales were linked across all available PISA waves from 2000 to 2018, they are basically comparable and thus suitable to be analyzed longitudinally (OECD, 2019b).

*Coding time* In the data sets, the time point of measurement was represented as the years starting from 2000 as the first measurement occasion coded as  $t_0 = 0$  and continuing in three-year intervals (2003 was coded as  $t_1 = 3$ , 2006 as  $t_2 = 6$ , and so on; see Fig. 4).

### **Data analysis procedure**

To tackle the guiding research questions (1) whether there was a linear trend in the PISA countries' mean reading scores, whether countries differed in the slope and direction of the trend, and (2) how persistent the effect of a potential deviation from the trend (a shock) would be, we first computed a series of models. For this purpose, we used the *ctsem* package (Version 3.4.3; Driver et al., 2017; Driver & Voelkle, 2018a, 2021) in R (R Core Team, 2021). All models were compared with respect to their fit. Because all models were nested in the CT-LCM-SR, multiple chi-square difference tests were performed. To address the question of whether there is a trend in PISA countries' mean reading literacy scores, we compared the CT-LCM-SR to a CT-AR(1) model, which is a CTM without trend component. To answer the question of whether there is also a dynamic process in the PISA data, we compared the CT-LCM-SR to a standard linear GCM (specified within the continuous-time framework), which is a linear trend model without a dynamic process. The dynamic process describes the effects of short-term influences (shocks) on countries mean reading literacy development. In addition, to answer whether there are significant differences in starting values and trends across PISA-countries, we compared fixed and random effects variants of all models (i.e., variances of the intercept and linear slope were freely estimated vs. fixed to zero).

We present an exemplary interpretation of the estimated parameters of the most relevant model, the CT-LCM-SR, focusing on the CT dynamic parameters. We also present and visualize DT parameters for different time-intervals derived from CT parameters and study the impact of temporary deviations from the trends on the development of countries' mean PISA reading scores to answer the second research question. Moreover, we visualize trends and dynamics of certain PISA countries as examples and show how predictions for future states can be made for various time-intervals.

### **Models**

In addition to the CT-LCM-SR, we also run a linear GCM and a standard CT-AR(1) model. We included fixed and random effects variants of all models, considering random effects for intercept and growth parameters but not for dynamic parameters (model specifications of all models can be found in the Additional file 1). These additional

models are nested in the CT-LCM-SR model. To obtain a *ctsem* version of a standard GCM from the CT-LCM-SR, we set the dynamic process component to zero, assuming a homoscedastic measurement error variance.<sup>6</sup> To obtain a standard CT-AR(1) model from the CT-LCM-SR, we fixed the growth parameter from the trend component to zero. To obtain the random effects variants of the models, we set the respective model parameters to be variant across countries (see the tutorial in the last part of this section). Pooling of estimates and variance components across the five data sets was carried out using the *mice* package (van Buuren et al., 2015).

**Fit Indices and Chi-Square tests** To be able to evaluate the relative model fits, we used *Akaike Information Criterion* (AIC; Akaike, 1973) and *Bayesian Information Criterion* (BIC; Schwartz, 1978) as well as the *Deviance* (twice the negative logarithm of the likelihood). For all three indices, lower values indicate better model fits. Furthermore, we run chi-square difference tests for nested models (e.g., Bollen, & Long, 1993). To pool the chi-square test statistics, the *miceadds* package (Robitzsch et al., 2017) was used.

**Missing data** Countries with less than five PISA participations between 2000 and 2018 were excluded from the analysis to ensure a sufficient data base for modeling trajectories. This procedure is in line with the analysis and presentation by the OECD (2019a). The remaining proportion of missing values in countries reading literacy scores was 9.4% and thus small.

Dealing with missing values is basically model-inherent in CTMs because the length of the time-interval between two consecutive measurement occasions is taken into account (for a detailed discussion of missing data treatment in CTMs, see Oud & Voelkle, 2014). When missing values are left in the data set, they can be addressed using, for example, maximum likelihood approaches (under the MAR assumption; Driver et al., 2017; Enders, 2010), as it was done in the present application.

#### ***ctsem* syntax for estimating the CT-LCM-SR**

In the following, we give a step-by-step tutorial of how to specify and run a CT-LCM-SR with the data from PISA using the *ctsem* package<sup>7</sup> (Driver & Voelkle, 2017; Driver et al., 2017) in R (R Core Team, 2021).

For continuous-time modeling with *ctsem*, two functions are crucial. The *ctModel* function is needed to set-up the required model, and the *ctStanFit* function is needed to fit the model to the data. Thus, in order to perform a CT-LCM-SR with the prepared PISA data, the first step was to specify the model. For setting up CTMs properly, the arguments of the *ctModel* function have to be understood. The following arguments are of key importance within the *ctModel* function (Driver & Voelkle, 2017; Hecht et al., 2019):

<sup>6</sup> The MANIFESTVAR matrix in *ctsem* was used to define the error term parameters of the GCM (see also Driver, 2020, and the specification in the Supplemental Material). Another more complicated alternative to estimate the standard GCM's homoscedastic error term (Fig. 3D) would be to constrain the variance of  $T0dyn_{SD}$  to  $q_{\Delta_1}^*$  (Fig. 3D). While the latter approach better captures the nested structure of GCM and CT-LCM-SR, it requires changes in the source code of *ctsem*, which is why using MANIFESTVAR to do this is much easier to implement in *ctsem*.

<sup>7</sup> To install the required *ctsem* package, we run the usual procedure in R (R Core Team, 2021):  

```
install.packages("ctsem")
library(ctsem)
```

```

ctModel(
  type,
  n.manifest,
  n.latent,
  manifestNames,
  latentNames,
  DRIFT,
  CINT,
  DIFFUSION,
  TOMEANS,
  TOVAR,
  MANIFESTMEANS,
  MANIFESTVAR,
  LAMBDA)

```

With the `type` argument, we chose the type of model we want to set-up. In *ctsem* the user has the option of `type="omx"` specifying a CTM interfacing to OpenMx (Neale et al., 2016), which is the original procedure of *ctsem* using frequentist estimation. Meanwhile, however, this part is mainly outsourced to an extra package *ctsemOMX* (Driver & Voelkle, 2021). Model type `"stanct"` interfaces to RStan<sup>8</sup> (Stan Development Team, 2020) and allows to choose between maximum likelihood, maximum a posteriori, or fully Bayesian estimation (Driver & Voelkle, 2017, 2021). In the present application we chose `type="stanct"` and maximum likelihood estimation for all models.

The argument `n.manifest` defines the number of variables to be analyzed in a given model. Because we wanted to analyze one single variable (countries' mean reading literacy scores), we set `n.manifest=1`. Furthermore, with `manifestNames`, the names of the variables to be analyzed as labelled in the respective data sets are defined. Reading scores were labelled "READ" in our data sets, so we set `manifestNames="READ"` (cf. Fig. 4). The argument `n.latent` determines the number of process components we need to analyze our variables under study. Because we needed two components for the CT-LCM-SR (one linear trend and one dynamic component) we set `n.latent=2`. With the `latentNames` argument, labels for the process components can be chosen. The number of names has to correspond with the `n.latent` argument. We chose `n.latent=c("lin", "dyn")`. The first process component should account for the trends in the data (linear trend) and the second should account for the dynamic process (fluctuating around the linear trends; Fig. 3A–C).

Next, the model matrices DRIFT, CINT, and DIFFUSION corresponding to Eq. 2 must be specified. The size of these three matrices corresponds to the number of process components (`n.latent`). Thus, in our application, the drift matrix **A** (`n.latent` × `n.latent`) was a 2 × 2 matrix. To model the linear trend component, the auto-effect of the first process had to be set to a negative value close to 0. In *ctsem* this specification is made automatically when a parameter on the main diagonal of the drift matrix is set to zero. In addition, the cross-effects between the two process components were also set to zero because trends and dynamics are independent in CT-LCM-SR. It was only the auto-effect of the

<sup>8</sup> DTMs are also possible via `type="standt"` (Driver & Voelkle, 2018a).

dynamic process that was freely estimated by setting  $\text{DRIFT} = \text{matrix}(c(0, 0, 0, "a"))$ . The diffusion matrix  $\mathbf{Q}$  looks very similar because in the CT-LCM-SR, we also only have a single continuous-time error process:  $\text{DIFFUSION} = \text{matrix}(c(0, 0, 0, "q"))$ . The continuous-time intercept vector  $\mathbf{b}$  ( $n.\text{latent} \times 1$ ) has a special role in the CT-LCM-SR specification in *ctsem* as it serves as a slope parameter in the trend process and therefore has to be freely estimated. As explained above, the continuous-time intercept was set to 0 via  $\text{CINT} = c("b", 0)$ .

Because the first measurement occasion cannot be regressed on a previous occasion, dynamic models must be initiated somehow. By default, *ctsem* uses a “predetermined” model with freely estimated parameters at the first time point (Driver et al., 2017). To model the linear trends of the CT-LCM-SR in *ctsem* (using Eq. 2), TOMEANS plays a special role and serves as the (freely estimated) intercept. Since the second process is (within-unit) centered in the CT-LCM-SR, it was restricted to be zero. So we set  $\text{TOMEANS} = c("int", 0)$ . Because the covariances of the initial time-points of both processes should be uncorrelated in the CT-LCM-SR, we set the  $2 \times 2$  matrix to  $\text{T0VAR} = \text{matrix}(c("intSD", 0, 0, "T0dynSD"))$ .

The two further arguments MANIFESTMEANS and MANIFESTVAR can be used to specify manifest components such as residuals or measurement errors (for more details, see Driver et al., 2017). In our application, we did not account for measurement error and set both to zero.

Finally, the LAMBDA (cf. Eq. 9) matrix relates the observed scores to the process components of the model (components are to be represented in the columns and the manifest variables in the rows):  $\text{LAMBDA} = \text{matrix}(c(1, 1))$ .<sup>9</sup> The complete specification is stored in a model object and reads as follows:

```
CT_LCM_SR <- ctModel(type='stanct',
  n.manifest=1,
  n.latent=2,
  manifestNames=c('READ'),
  latentNames=c('lin', 'dyn'),
  DRIFT=matrix(c(0, 0, 0, 'a'), nrow=2, ncol=2),
  DIFFUSION=matrix(c(0, 0, 0, 'q'), nrow=2, ncol=2),
  T0VAR=matrix(c('intSD', 0, 0, 'T0dynSD'), nrow=2, ncol=2),
  CINT=matrix(c('b', 0), nrow=2, ncol=1),
  TOMEANS=matrix(c('int', 0), nrow=2, ncol=1),
  LAMBDA=matrix(c(1, 1), nrow=1, ncol=2 ),
  MANIFESTMEANS=c(0),
  MANIFESTVAR=matrix(c(0), nrow=1, ncol=1 ) )
```

After model specification, we can print and check the model object with the command `head(CT_LCM_SR$pars, 20)`. In addition, some further modifications can be made, such as fixing individual parameters to a certain value or specifying random effects (Driver & Voelkle, 2018b). In the current application, we tested if countries differed in terms of their initial level at the first PISA wave in 2000 (random intercept) as well as in terms

<sup>9</sup> This specification is only possible in the case of linear trends. We provide an example code in the supplemental material how to include quadratic trends in *ctsem* (and examples can also be found on GitHub; Driver, 2022). In the case of quadratic effects, the PARS argument has to be used and the LAMBDA matrix must be employed in a different way.

**Table 1** Pooled Means and Standard Deviations of PISA Reading Literacy Scores Aggregated at the Country Level

Time points	PISA waves	M	SD	Range	N
T1	2000	474.8	53.9	322.0–549.1	42
T2	2003	482.6	41.0	374.1–543.8	40
T3	2006	466.0	51.3	311.5–556.8	53
T4	2009	470.5	43.9	370.1–540.0	56
T5	2012	475.3	45.0	384.6–544.9	56
T6	2015	473.4	42.2	359.1–527.7	54
T7	2018	470.9	40.9	371.3–525.6	54
-	Overall	473.0	45.4	311.5–556.8	355 <sup>a</sup>

The descriptive statistics were calculated based on the countries' mean scores that form the data basis of the longitudinal analysis in the present study. For the 2015 and 2018 waves, the achievement data in PISA were represented by 10 plausible values instead of 5. Due to the longitudinal modeling, only the first 5 plausible values were considered in the present study

<sup>a</sup> The overall *N* represents the number of data points in the data set (number of countries in the sample \* the respective number of repeated measurements). A total of 56 countries were included in the sample the inclusion criterion being at least 5 PISA participations between 2000 and 2018

of direction and slope of the trends (random slope) and specified random effects for  $t0m$  (random intercept) and  $b\_lin$  (random slope): `CT_LCM_SR$vars[c(1, 7),]$indvarying <- TRUE`. Any parameter that should vary across units (in this case parameter one, which is the intercept, and parameter seven, which is the slope) must be set to `indvarying = TRUE` in the model object.

Once all parameters were properly specified, the model could be fitted to the PISA data. For this purpose, the `ctStanFit` function was used. For maximum likelihood estimation we set the following additional arguments (see Driver & Voelkle, 2021): `CT_LCM_SR_fit <- ctStanFit(datalong = PISA_PV1, ctstanmodel = CT_LCM_SR, optimize = TRUE, nopriors = TRUE)`.

### Descriptive statistics

For the sample of  $N=56$  countries (level-2 units) included in the analysis the descriptive statistics of the mean reading literacy scores for each of the seven waves are reported in Table 1. In the present sample, the overall mean across the 56 countries and all seven waves was 473 PISA points ( $SD=45.4$ ), pooled over the five data sets.

### Model comparisons and fit statistics

Model comparisons and fit statistics provided first insights regarding the research questions. In Table 2, the deviance, AIC and BIC of the six models are presented. The random effects models outperformed the fixed effects variants in all cases. This implied considerable differences in initial levels and trends across countries' trajectories of mean PISA reading literacy scores. Furthermore, all three information criteria showed the smallest values for the CT-LCM-SR with random effects, indicating the best model fit of all six models. In addition, we also performed chi-square tests for nested models. In all cases the chi-square test was significant at a two-tailed significance-level of 0.05, also indicating the best model fit for the random effects CT-LCM-SR.<sup>10</sup> Thus, the comparisons of

<sup>10</sup> Moreover, we tested a GCM with additional quadratic effects. The quadratic effects did not have a statistically significant contribution and were therefore not considered in further analysis (e.g., Voelkle, 2008).

**Table 2** Comparison of fit indices in models fitted to PISA data

Models	Number Est. Parameters	Deviance	AIC	BIC
Fixed effects models				
GCM (FI-FG)	3	3715.1	3721.1	3733.1
CT-AR(1) model (FI)	4	3043.7	3051.7	3067.6
CT-LCM-SR (FI-FG)	5	3009.9	3019.9	3039.8
Random effects models				
GCM (RI-RG)	5	3007.8	3019.8	3043.6
CT-AR(1) model (RI)	6	3005.0	3015.0	3034.8
CT-LCM-SR (RI-RG) <sup>a</sup>	8	2984.9	3000.9	3032.6

AIC Akaike information criterion, BIC Bayesian information criterion, FI fixed intercept, FG fixed growth parameter, RI random intercept, RG random growth parameter. The CT-AR(1) has just intercepts, but no growth parameters

<sup>a</sup> Details on this model are provided in Table 3

**Table 3** Parameter Estimates of the CT-LCM-SR for PISA Countries' Mean Reading Literacy Scores Development

Parameter (labels <sup>a</sup> )	Estimate	SE	t	p
Fixed effects				
Intercept (int)	464.80	7.37	63.034	< .001
Linear growth (b)	0.43	0.20	2.148	.037
CT auto-effect (a)	− 0.39	0.13	− 3.099	.003
Variance components (in SD metric)				
Intercept SD (int_SD)	53.50	5.57	9.597	< .001
Growth SD (b_SD)	1.06	0.32	3.360	.002
Initial residual SD (T0dynSD)	9.95	0.86	11.527	< .001
Diffusion SD (q)	18.58	2.99	6.206	< .001

The random effects CT-LCM-SR contains an eighth parameter, the intercept-growth correlation, which was estimated to be − 0.61 ( $SE=0.13$ ,  $p<.001$ )

<sup>a</sup> The labels refer to the parameter labels used in Eq. 10 and 11 as well as in the presentation of the R code (p. 23) and Fig. 2

these models provided first support for the existence of trends, differences in trends and initial levels across countries, and a dynamic process.

### Results of the CT-LCM-SR for the PISA countries' reading literacy scores

The parameter estimates of the random effects CT-LCM-SR, pooled over the analyses of all five data sets, are shown in Table 3 (parameter estimates of the CT-AR(1) and the GCM can be found in the Table 5 in Appendix). The CT-LCM-SR yielded eight estimated parameters, which were all statistically significantly different from zero. This supported the use of the CT-LCM-SR, which accounts for dynamics and trends in the data. Additionally, this also supported the use of the random effects CT-LCM-SR, which allows for differences in terms of initial values and slopes across countries.

### Examining trends in PISA countries' mean reading literacy scores

Five parameters of the CT-LCM-SR were related to the trends, namely the mean intercept, the intercept standard deviation ( $SD$ ), the growth rate, the growth rate  $SD$ , and the intercept-growth correlation. The interpretation of these five parameters corresponds

with the standard parameters of a GCM with random effects. The estimated mean intercept across all countries was 464.8 PISA points ( $SE = 7.37$ ), and the corresponding intercept standard deviation was 53.50 ( $SE = 5.57$ ). The estimated growth rate was 0.43 ( $SE = 0.20$ ) per year, which implies an average increase of 1.29 PISA points per measurement occasion ( $3 * 0.43$ ) across all countries. The corresponding standard deviation was  $SD = 1.06$  ( $SE = 0.32$ ). Thus, for instance, for a country with a growth rate of one  $SD$  below the average, a negative trajectory is expected and thus a loss of  $-0.63$  ( $0.43 - 1.06$ ) PISA points per year on average (and  $-1.89$  for the three-year interval between adjacent PISA waves). The estimated correlation of  $-0.61$  ( $SE = 0.13$ ) between intercept and growth rate indicated that countries with lower mean scores at the initial PISA wave in 2000 tended to make larger gains in mean reading literacy scores on the follow-up surveys than countries with higher mean scores in 2000.

Thus, with respect to the first research question, we found a significant overall increase of 1.29 PISA points per wave or 0.43 PISA points per year in the sample of 56 PISA countries. However, countries differed significantly in the slope and even the direction of the trend (e.g., a country that is one  $SD$  below the average is expected to have a negative trend). In addition, we found that countries also differ substantially in terms of their initial values at the first measurement occasion in 2000, and that countries with lower initial scores tend to achieve larger gains in their mean reading literacy scores across subsequent PISA waves.

### **Examining the dynamics in the development of PISA countries' mean reading literacy scores**

While the trend parameters of the CT-LCM-SR describe static change processes (Fig. 3A), the parameters of the dynamic process component describe fluctuations (Fig. 3B). These fluctuations were of substantive interest with respect to the second research question, which concerned the question of how long deviations from the trends (shocks) influence a country's development of mean PISA reading literacy performance. As mentioned above, such deviations can be caused by short-term support programs, or temporary investments in the education system as well as a financial crisis or a pandemic.

The five parameters related to the first process component provide information about the linear trends in the data and center the residuals that enter the dynamic process at the within-country level (Fig. 3A–C). Thus, the dynamic process can be considered as *detrended* as the growth component accounts for the linear trends in the data (e.g., Usami et al., 2019). The remaining three parameters of the CT-LCM-SR, namely the auto-effect, the diffusion variance, and the initial residual variance component, are related to the dynamic process. The initial variance component represents the residual variance of the first measurement occasion entering the dynamic process and, thus, yields the initial (predetermined) values of the dynamic process in the CT-LCM-SR. The diffusion variance represents the CT error variance, which is the part of the variance of the process that is completely random (Oud & Jansen, 2000). The CT auto-effect can be used to study the time course of the impact of temporary deviations (also referred to as shocks) on the process under study and was therefore the relevant parameter to answer the second research question.

**Table 4** Discrete-time parameters for autoregression (AR) and process error for varying time-intervals derived from underlying continuous-time parameters

Time-interval (years)	DT AR effects	95% CI for AR effects	DT process error (SD)
1	.681	[.841; .521]	6.91
2	.471	[.684; .257]	10.08
3	.329	[.546; .112]	11.53
4	.233	[.433; .033]	12.19
5	.167	[.342; – .009]	12.50
6	.121	[.271; – .029]	12.64
9	.048	[.136; – .039]	12.75

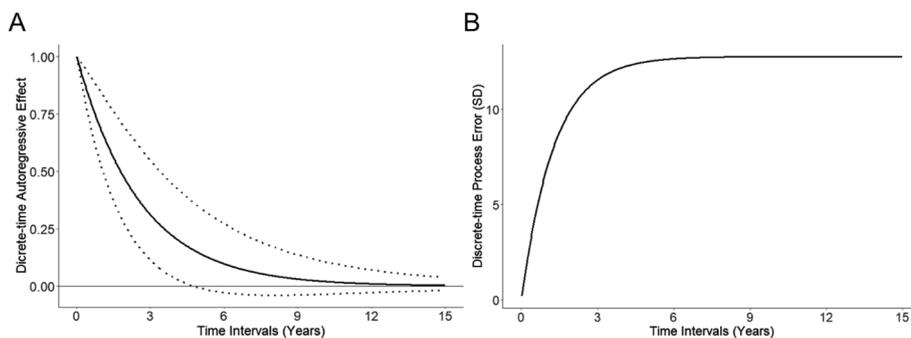
CI confidence interval. A three-year interval represents the time between two PISA surveys. Values for the time-intervals 1, 2, 4 and 5 are interpolations, because there is no data for such time-intervals in the data sets

The interpretation of CT auto-effects is different from autoregressive effects of DTMs. Typical for stable or mean-reverting dynamical processes is the negative value of the auto-effect (Ryan et al., 2018). Translated to DTMs, negative CT auto-effects mean that the corresponding DT autoregressive parameters are between 0 and 1. Furthermore, DT autoregressive effects get smaller for longer time-intervals between consecutive measurement occasions in such stable processes. The negative CT auto-effect refers to the fact that if the value of a time series at time  $t$  takes a position far from the process mean the process subsequently tends to return to the process mean with the opposite sign to the deviation (assuming that there are no further shocks afterwards). Furthermore, a negative CT auto-effect indicates that the more distant a deviation from the process mean at a time  $t$ , the greater is the speed with which it tends to return to the process mean (assuming the same underlying auto-effect). Likewise, the more negative the auto-effect, the faster the process tends to return to the process mean after a deviation from the trend (for the same time-interval).

Because DT autoregressive effects for various time-intervals can be derived from the estimated CT auto-effect, this parameter is useful to answer the question of how persistently deviations from the trend affected countries' development of mean reading literacy achievement. The CT auto-effect was estimated to be  $-0.39$  ( $SE=0.13$ ). It is important to note that the auto-effect was the same for all countries of the sample in this model (no random effects), unlike the (country-specific) parameters of the trends. To better understand the dynamic effects and find an answer to the second research question, DT autoregressive parameters can be derived from the CT auto-effect for various intervals.<sup>11</sup> DT autoregressive effects can be interpreted as “the amount of within-person carry over effect” from one measurement occasion to the next (Hamaker et al., 2015, p. 10 l. 13<sup>12</sup>). CT modeling additionally provides information on the dependence of the DT autoregressive effects on the time interval length between measurement occasions.

<sup>11</sup> In multivariate models the same can be done for all lagged parameters including cross-effects (e.g., Ryan et al., 2018).

<sup>12</sup> Recognize the difference in interpretation between fixed and random effects models discussed by Hamaker et al. (2015) in the same article.



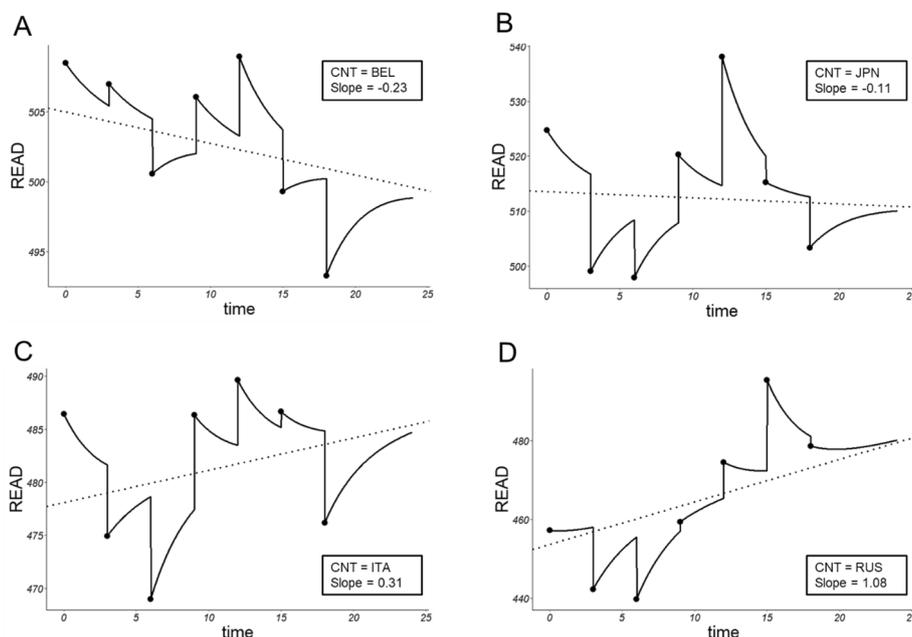
**Fig. 5** Discrete-time autoregression (A) and Error standard deviation (B) as Functions of Underlying Continuous-Time Parameters and the Length of Time-Intervals Between Measurement Occasions. Dotted lines represent the 95% confidence interval of the discrete-time autoregressive effects corresponding to Table 4 column 3

### Deriving discrete-time parameters as functions of underlying continuous-time parameters

Using Eq. 3, we can calculate DT parameters for arbitrary time-intervals from the estimates of the underlying CT parameters. In Table 4 we present a range of DT autoregressive coefficients and the related confidence intervals. We can do the same for the diffusion variance and derive DT error variances, using Eq. 5 (Table 4). For longer time-intervals, the autoregressive coefficient becomes smaller, whereas the error variance becomes larger (Oravecz et al., 2018). DT autoregressive effects are usually interpreted in the literature as indicating the stability or the inertia of the construct under study (e.g., Hamaker et al., 2015). Thus, they provide information on how much of the deviation at time  $t_i$  is carried over to  $t_{i+1}$  (relative to the process mean) on average. To address the second research question, the time course of the autoregressive effects and the associated confidence intervals can be used (Table 4). The confidence interval of the autoregressive effect for the five-year interval is the first that includes zero, that is, there is no evidence for a population autoregressive effect larger than zero. An autoregressive effect of zero for a time interval of five-years or longer implies that after this period earlier deviations from the trend have no predictive value. To visualize this effect, DT coefficients can be represented as continuous functions of the (time independent) CT parameters and the length of time-interval. As can be seen in Fig. 5, for increasing length of the time-interval between measurement occasions, the DT autoregressive effect approaches zero (Fig. 5A), while the error standard deviation approaches about 13 (Fig. 5B).

### Representing countries-specific expected trajectories and making predictions for future states

Unlike DTMs, in which processes can only be represented with respect to the investigated interval (see Fig. 1B), CTMs such as the CT-LCM-SR allow to represent the time course of developmental trajectories. Combining trends and dynamics, Fig. 6 shows the country-specific trajectories of four PISA countries in the sample. One major advantage of CTMs like the CT-LCM-SR is that predictions can be made with respect to any time-interval. If PISA reading literacy scores were analyzed using a DTM, the predictions of future states would be limited to the three-year intervals (and their multiples), because



**Fig. 6** Model implied expected trajectories given past observations of four PISA countries' mean reading literacy score development. *CNT* = country code, Slope = country specific growth rate (dotted lines). The solid black line represents the expected model implied trajectory given past observations for a respective country. The predictions are a combination of both process components (linear trend and dynamics; cf. Eq. 8, 9). The slope parameter estimates are obtained from the trend component of the CT-LCM-SR and are represented by the dotted lines. The deviations from the trend and the subsequent carry-over effects are described by the dynamic component of the CT-LCM-SR

of the time-interval dependency of DTMs. Moreover, unlike predictions made by a simple GCM, the CT-LCM-SR uses the information of the residuals, which are fluctuating around the static trends (the grey dotted lines in Fig. 6). Predictions of a GCM would always lie exactly on the static trend and would not take into account the dynamics of the residuals. Figure 6 illustrates the mean reverting dynamic process fluctuating around the country-specific static trends. The size of the autoregressive effect decreases as the length of the time-intervals increases, representing the impact of deviations (shocks) on later states over time. The autoregressive effect reaches zero after a certain length of the time-interval (see also Fig. 5A) indicating that there is no further influence of the previous deviation. For long time-intervals the autoregressive effect is zero and the best prediction lies on the linear trend. Thus, the point predictions return to the static linear trend process after a certain interval (Fig. 6).

## Discussion

The aim of the present study was to discuss advantages of CTMs, especially the newly proposed CT-LCM-SR, for longitudinal data analysis in the education sciences and to provide the reader with an easy-to-follow introduction to specifying and interpreting CTMs with the help of an illustrative example using PISA data.

In the first section, we discussed that CTMs are often better suited to capture the continuous nature and development of constructs in educational research than DTMs. Using CTMs, we can thus better align our statistical models with our theories of change

that are the subject of educational research. This includes not only variables at the level of individuals but also variables that describe characteristics of entire education systems as in the exemplary application with PISA data. In addition, we outlined other advantages of CTMs over DTMs associated with the time-interval dependency of DTMs, such as limitations in cross-study comparisons, limitations in the procedures of data collection, and the lack of relevant information regarding the dynamic processes.

Based on this, we developed the CT-LCM-SR, a CTM variant that is suitable for a range of typical longitudinal research questions. The CT-LCM-SR takes into account both the dynamic process as well as trends in the data. Trends are often of substantive interest in the education sciences and are also relevant for dynamic models because unaccounted trends can lead to biases in parameter estimates (Asparouhov et al., 2018; Núñez-Regueiro et al., 2021; Walls & Schäfer, 2006). The process' dynamics can provide information about the stability or variability of constructs, their self-related structure, and about reciprocal effects of multiple variables in multivariate models. In addition to the advantage that (1) the CT-LCM-SR provides information on both trends and dynamics, the use of the CT-LCM has further advantages. (2) Trends and dynamic processes are clearly separated in the CT-LCM-SR, meaning that trends and dynamics cannot influence or "compete" with each other. (3) The interpretation of the parameters follows the interpretation of CTMs and GCMs. (4) All parameters of the CT-LCM-SR describe the process under study independently of the length of time-intervals in a given study (Delsing & Oud, 2008). Finally, (5) both the standard linear GCM and the CT-AR(1) model are nested within the CT-LCM-SR, and thus, these models can be tested against the CT-LCM-SR using chi-squared difference testing. In addition, with the CT-LCM-SR, we can take advantage of all the other benefits associated with CTMs.

In the second part, we demonstrated the application of the CT-LCM-SR with the R package *ctsem* (Driver & Voelkle, 2017, 2021). The analysis was guided by two exemplary research questions. First, we wanted to examine whether there is a significant trend in the mean PISA reading literacy scores of countries over the period from 2000 to 2018, and whether countries differ in the slope and direction of trends. Second, we aimed to study the stability of education systems when deviations from the trend occur. Examining the persistence of deviations is relevant to study the system's resilience against disturbances caused by different influences.

Regarding the first research question, we found an overall increase in PISA countries' mean reading literacy scores. However, countries differed considerably in the slope and the direction of trends. Moreover, countries also differed with respect to the initial levels at the first measurement occasion in 2000. We found that countries with lower scores at the first measurement occasion tended to achieve larger gains in mean reading literacy scores on the follow-up surveys than countries with higher initial scores.

Regarding the second research question, we found a significant continuous-time dynamic process of fluctuations around the trends. From this dynamic process it can be inferred how long temporary deviations from the trends are associated with later states of the variable(s) under study and how much of the deviation is carried over to future

states depending on the time that has passed. Using the CT auto-effect, we derived DT autoregressive effects for various time-intervals and found that carry-over effects (represented by DT autoregressions) for five-year or longer time-intervals were not significantly different from zero. This means that both positive and negative influences on countries' mean reading literacy scores that led to temporary deviations from the trends disappeared on average after about five years.

### Limitations and future directions

By introducing the CT-LCM-SR, the current article suggests and illustrates an advanced method with several benefits and potentials for longitudinal studies in the education sciences. However, the presented work is not without limitations.

First, it is necessary to consider that the application presented is a simplified example chosen in order to provide an easy introduction to the potential benefits, practical modeling, and interpretation of CTMs. Therefore, the substantive value of the results is limited. Only a single variable was analyzed, the mean PISA reading scores of countries. For some applications, such as predictions of future states, univariate models can already provide useful information (e.g., Bulteel et al., 2018). Of course, however, research questions are often more complex. In the present application, for example, it is only possible to speculate on why temporary deviations from trends occur, as we are studying *shocks* of unknown cause. In practical applications, it might be of interest to examine the impact of a specific intervention (or crises). This requires the inclusion of an input that occurs at specific points in time in the model (see Driver & Voelkle, 2018b for the implementation of such models in *ctsem*). Such models make it possible to examine, for instance, when the input reaches the maximum effect on the process under study and after what period of time the effect of an input vanishes.

Moreover, the dynamic interplay between two or more parallel processes might be of researchers' interest (see, e.g., Jindra et al., under review, to be published in the same special issue applying a bivariate CT-LCM-SR). Such models allow, for instance, to explore the time interval in which one process has a maximum effect on the other and vice versa (e.g., Hecht & Zitzmann, 2021b). A typical question for educational research might be after how much time a change in a student's achievement does develop its maximum effect on his or her self-concept. R code for studying the reciprocal effects of two variables over time in a bivariate CT-LCM-SR is provided in the Additional file 1.

In addition, the use of covariates plays a crucial role in practical research, including the study of trends in countries' mean PISA scores, for which so-called *adjusted trends* have been reported meanwhile (OECD, 2014). Such more complex CTMs with predictors and covariates can be implemented with *ctsem* (e.g., Driver & Voelkle, 2018b).

It is also important to keep in mind, that development at the country-level was examined, not at the student level. While short-term interventions might not change the performance of the entire education system permanently, they might have effects on the students' cohort. However, statements concerning developmental trajectories of students require repeated measurements of the same students. A first application of a

bivariate CT-LCM-SR on the student-level is provided by Jindra et al. (under review) to be published in the same special issue.

Another possibility in *ctsem* is specifying measurement models (e.g., Hecht et al., 2019). In the presented application with PISA data, a typical three-step procedure was chosen to address measurement error (cf. OECD, 2009). We first calculated aggregate country means for each of the five plausible values and for all seven available PISA waves and created five data sets (so that all plausible values could be considered). These data sets were then used to conduct all analyses separately. Finally, the obtained results were pooled across the five separate analyses (Rubin, 2004).

The CTMs presented must also be considered with some limitations and model assumptions in mind. In the present article, we have restricted ourselves to univariate stable dynamic processes, also called mean-reverting processes (Ryan et al., 2018). That implies that the corresponding DT autoregressive effects are between 0 and 1 for any time-interval. In principle, however, CTMs for explosive processes, meaning DT autoregressions greater than 1 (e.g., Driver & Voelkle, 2021), or processes with negative DT autoregressions, may be of interest as well (Fisher, 2001). In this article, we furthermore restricted ourselves to the representation of first-order CTMs, which should be suitable for representing many processes studied in educational research (Sivo & Fan, 2008). However, modeling higher-order models is also possible and might be useful in certain applications (e.g., Lüdtke & Robitzsch, 2021; Oud, 2010; Oud et al., 2018). Furthermore, in the models presented, the dynamic effects were assumed to be invariant over the period studied. This means, for example, that the CT auto-effect does not change and the DT autoregressive effects change only as a function of the length of the time-interval but not with time itself. However, more complex models with time-varying dynamic effects are also possible (e.g., Driver & Voelkle, 2021; Oud & Jansen, 2000). In addition, in the application presented, the dynamic effects were assumed to be the same across all observation units. Random effects were only included in the trend components but not in the dynamic process. This is a limitation that can be overcome within *ctsem* because fully hierarchical models can be specified in which all parameters vary across units (Driver & Voelkle, 2018a). Future simulation studies are needed to further investigate the estimation performance of CTMs in general and the CT-LCM-SR in particular with respect to different influencing factors, such as sample size and length of individual time series (for first sample-size recommendations and a discussion of the compensating effects of sample size  $N$  and length of time series  $T$  in CTMs, see Hecht & Zitzmann, 2021b).

In the CT-LCM-SR, we proposed to consider linear trends. The use of linear trends is not without criticism and can usually only be justified for the duration of specific time periods, as they always tend toward infinity in the long run (e.g., Oud, 2010). However, linear trends have a clear and well-known interpretation, which can be an advantage. Nevertheless, other types of trends might be considered and in future research, continuous-time models that incorporate other than linear trends could be developed or further refined (e.g., Oud, 2010; Voelkle & Oud, 2015). For example, in the PISA example, we also tested a GCM with nonlinear (quadratic) components, which did however not reach

significance. We provide example code for setting up GCMs and the CT-LCM-SR with quadratic trend components in *ctsem* (similar code can be found on the GitHub page of *ctsem*; Driver, 2022).

In the CT-LCM-SR as applied in the present article, it remains uncertain how a given shock affects the trend. Models, which allow a shock to influence the direction and shape of the trend, are also possible (Driver & Voelkle, 2018b).

In educational research, specific data situations may occur which demand appropriate analysis strategies like resampling and missing data imputation techniques (e.g., Weirich et al., 2021). Future research should generate guidance on how to address such educational research specific challenges when applying CTMs. Furthermore, estimation of CTMs might become very time-consuming, especially in educational large-scale assessment contexts. One promising optimization approach to reduce run times for continuous-time models has been proposed by Hecht and Zitzmann (2020) based on the work of Hecht et al. (2020).

Finally, although CTMs provide information about the development of dynamic effects, interpolations to time-intervals that have not been investigated should be considered with caution. While it may be useful to explore intervals of maximum effects based on CTMs (Hecht & Zitzmann, 2021a), it is not advisable to draw conclusions about change processes that occur at much smaller intervals, for example days or hours, based on annual surveys, for instance. Thus, even when using CTMs, the underlying theory of change should still guide decisions about the design of longitudinal studies. However, equal time-intervals between subjects and measurement occasions are not necessary when using CTMs, which allows for much greater flexibility in data collection than when using DTMs (see Voelkle & Oud, 2013, for an argument why it can even be beneficial to choose time-intervals of different length between individuals and measurement occasions).

## Conclusion

In the present article, we presented continuous-time modeling as a suitable approach to align statistical models for longitudinal data analysis with theories of change, especially to represent the continuous nature of variables and observation units. It is essential to align theories and statistical models to obtain valid answers to the research questions. In addition, the use of CTMs comes with other considerable advantages which were mentioned. Based on this, we proposed the CT-LCM-SR, which is a CTM variant and a suitable model to address many research questions in the education sciences. The CT-LCM-SR yields parameters, which are independent of specific time-intervals, thus represents developmental processes on a continuous-time scale. We illustrated the implementation and interpretation of the CT-LCM-SR with PISA data and, thus, gave the reader an introduction to the use of CTMs and the advantages and functionality of the newly proposed CT-LCM-SR.

## Appendix

See Table 5.

**Table 5** Parameter Estimates of the CT-AR(1) model, the Standard GCM and the CT-LCM-SR

Parameters (labels) <sup>a</sup>	Process <sup>b</sup>			GCM			CT-LCM-SR		
	Estimate	SE	p	Estimate	SE	p	Estimate	SE	p
Fixed effects									
Intercept (int)	469.73	6.02	<.001	464.69	7.44	<.001	464.80	7.37	<.001
Linear growth (b)	–	–	–	0.49	.45	.018	0.43	0.20	.037
CT auto-effect (a)	–.19	0.05	<.001	–	–	–	–0.39	0.13	.003
Variance components									
Intercept SD (int_SD)	42.91	4.47	<.001	54.56	5.52	<.001	53.50	5.57	<.001
Growth SD (b_SD)	–	–	–	1.30	0.43	<.001	1.06	0.32	.002
Initial residual SD (T0dynSD)	26.25	4.42	<.001	–	–	–	18.58	2.99	<.001
Residual (m_err)/Diffusion SD (q)	8.80	0.57	<.001	10.65	0.70	<.001	9.95	0.86	<.001
Covariances									
Intercept-Growth (Cor_T0M_b)	–	–	–	–.55	.23	<.001	–.61	.13	<.001

The parameter estimates of the CT-LCM-SR were already represented in Table 3 and are repeated here to facilitate comparisons

<sup>a</sup>The labels refer to the parameter labels used in Eq. 10 and 11 as well as in the presentation of the R code (p. 23) and Fig. 2

<sup>b</sup>This column indicates, which parameter belongs to which process component

**Abbreviations**

AIC	Akaike Information Criterion
ALT	Autoregressive latent trajectory (model)
AR	
(model)	Autoregressive model
BIC	Bayesian Information Criterion
CALT	Continuous-time autoregressive latent trajectory (model)
CT-AR	
(model)	Continuous-time autoregressive (model)
CT-LCM-SR	Continuous-time latent curve model with structured residuals
CT-VAR	
(model)	Continuous-time vector autoregressive (model)
CTM	Continuous-time model
DTM	Discrete-time model
GCM	Growth-curve model
LCM-SR	(Discrete-time) latent curve model with structured residuals
VAR	
(model)	Vector autoregressive model

**Supplementary Information**

The online version contains supplementary material available at <https://doi.org/10.1186/s40536-022-00126-8>.

**Additional file 1.** Supplementary R Code.

**Acknowledgements**

Not applicable.

**Author contributions**

The authors declare the following contributions (as defined by <http://credit.niso.org> (accessed on 16 January 2022)) to this article: JL: conceptualization, data curation, formal analysis, investigation, methodology, project administration, software, supervision, validation, visualization, writing—original draft, writing—review and editing. SZ: conceptualization, methodology, project administration, supervision, validation, writing—review and editing. MV: conceptualization, methodology, supervision, validation, writing—review and editing. MH: conceptualization, formal analysis, methodology, project administration, software, supervision, validation, writing—review and editing.

**Funding**

No sources of funding to declare.

**Availability of data and materials**

The data analyzed in this article are available at the PISA website: <https://www.oecd.org/pisa/data/>. The R code for all models run for the present analyses and an additional script for a bivariate CT-LCM-SR are available at the following link : <https://doi.org/10.6084/m9.figshare.19850926>

**Declarations****Ethics approval and consent to participate**

Not applicable.

**Consent for publication**

All authors provide consent for publications.

**Competing interests**

We have no conflict of interest to disclose.

Received: 3 February 2022 Accepted: 24 June 2022

Published online: 11 July 2022

**References**

- Asparouhov, T., Hamaker, E. L., & Muthén, B. (2018). Dynamic structural equation models. *Structural Equation Modeling*, 25(3), 359–388. <https://doi.org/10.1080/10705511.2017.1406803>
- Baltes, P. B., & Nesselroade, J. R. (1979). History and rationale of longitudinal research. In J. R. Nesselroade & P. B. Baltes (Eds.), *Longitudinal research in the study of behavior and development* (pp. 1–39). Academic Press.
- Bianconcini, S., & Bollen, K. A. (2018). The latent variable-autoregressive latent trajectory model: A general framework for longitudinal data analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, 25(5), 791–808. <https://doi.org/10.1080/10705511.2018.1426467>
- Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: A structural equation perspective* (Vol. 467). Wiley. <https://doi.org/10.1037/0022-0663.92.4.703>

- Bollen, K. A., & Long, S. J. (Eds.). (1993). *Testing structural equation models*. Sage.
- Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: Forecasting and control*. Wiley.
- Bulteel, K., Mestdagh, M., Tuerlinckx, F., & Ceulemans, E. (2018). VAR(1) based models do not always outpredict AR(1) models in typical psychological applications. *Psychological Methods*, 23(4), 740–756. <https://doi.org/10.1037/met0000178>
- Caro, D. H., & Biecek, P. (2017). intsvy: An R package for analyzing international large-scale assessment data. *Journal of Statistical Software*, 81(7), 1–44. <https://doi.org/10.18637/jss.v081.i07>
- Collins, L. M. (2006). Analysis of longitudinal data: The integration of theoretical model, temporal design, and statistical model. *Annual Review of Psychology*, 57, 505–528. <https://doi.org/10.1146/annurev.psych.57.102904.190146>
- Corp, I. B. M. (2020). *IBM SPSS statistics for windows, version 27.0*. IBM Corp.
- Curran, P. J., Howard, A. L., Bainter, S. A., Lane, S. T., & McGinley, J. S. (2014). The separation of between-person and within-person components of individual change over time: A latent curve model with structured residuals. *Journal of Consulting and Clinical Psychology*, 82(5), 879–894. <https://doi.org/10.1037/a0035297>
- de Haan-Rietdijk, S., Voelkle, M. C., Keijsers, L., & Hamaker, E. L. (2017). Discrete- vs. Continuous-time modeling of unequally spaced experience sampling method data. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2017.01849>
- Deboeck, P. R., & Boker, S. M. (2015). Analysis of dynamical systems: the modeling of change and variability. In S. J. Henly (Ed.), *Routledge international handbook of advanced quantitative methods in nursing research* (pp. 170–186). Routledge/Taylor & Francis.
- Deboeck, P. R., & Boulton, A. J. (2016). Integration of stochastic differential equations using structural equation modeling: A method to facilitate model fitting and pedagogy. *Structural Equation Modeling*, 23(6), 888–903. <https://doi.org/10.1080/10705511.2016.1218763>
- Deboeck, P. R., & Preacher, K. J. (2016). No need to be discrete: A method for continuous time mediation analysis. *Structural Equation Modeling*, 23(1), 61–75. <https://doi.org/10.1080/10705511.2014.973960>
- Deboeck, P. R., Nicholson, J., Kourou, C., Little, T. D., & Garber, J. (2015). Integrating developmental theory and methodology: Using derivatives to articulate change theories, models, and inferences. *Applied Developmental Science*, 19(4), 217–231. <https://doi.org/10.1080/10888691.2015.1021924>
- Delsing, M. J., & Oud, J. H. (2008). Analyzing reciprocal relationships by means of the continuous-time autoregressive latent trajectory model. *Statistica Neerlandica*, 62(1), 58–82. <https://doi.org/10.1111/j.1467-9574.2007.00386.x>
- Dormann, C., Guthier, C., & Cortina, J. M. (2020). Introducing continuous time meta-analysis (CoTiMA). *Organizational Research Methods*, 23(4), 620–650. <https://doi.org/10.1177/1094428119847277>
- Driver, C. C., & Voelkle, M. C. (2018b). Understanding the time course of interventions with continuous time dynamic models. In K. Van Montfort, J. H. L. Oud, & M. C. Voelkle (Eds.), *Continuous time modeling in the behavioral and related sciences* (pp. 79–109). Springer International. [https://doi.org/10.1007/978-3-319-77219-6\\_4](https://doi.org/10.1007/978-3-319-77219-6_4)
- Driver, C. C., & Voelkle, M. C. (2018a). Hierarchical Bayesian continuous time dynamic modeling. *Psychological Methods*, 23(4), 774–799. <https://doi.org/10.1037/met0000168>
- Driver, C. C., Oud, J. H., & Voelkle, M. C. (2017). Continuous time structural equation modeling with R package ctsem. *Journal of Statistical Software*. <https://doi.org/10.18637/jss.v077.i05>
- Enders, C. K. (2010). *Applied missing data analysis*. Guilford Press.
- Gollob, H. F., & Reichardt, C. S. (1987). Taking account of time lags in causal models. *Child Development*, 58(1), 80–92. <https://doi.org/10.2307/1130293>
- Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*. <https://doi.org/10.2307/1912791>
- Hamaker, E. L. (2005). Conditions for the equivalence of the autoregressive latent trajectory model and a latent growth curve model with autoregressive disturbances. *Sociological Methods & Research*, 33(3), 404–416. <https://doi.org/10.1177/0049124104270220>
- Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods*, 20(1), 102. <https://doi.org/10.1037/a0038889>
- Hecht, M., & Voelkle, M. C. (2021). Continuous-time modeling in prevention research: An illustration. *International Journal of Behavioral Development*, 45(1), 19–27. <https://doi.org/10.1177/0165025419885026>
- Hecht, M., & Zitzmann, S. (2020). A computationally more efficient Bayesian approach for estimating continuous-time models. *Structural Equation Modeling: A Multidisciplinary Journal*, 27, 829–840. <https://doi.org/10.1080/10705511.2020.1719107>
- Hecht, M., & Zitzmann, S. (2021a). Exploring the unfolding of dynamic effects with continuous-time models: Recommendations concerning statistical power to detect peak cross-lagged effects. *Structural Equation Modeling: A Multidisciplinary Journal*, 28(6), 894–902. <https://doi.org/10.1080/10705511.2021.1914627>
- Hecht, M., & Zitzmann, S. (2021b). Sample size recommendations for continuous-time models: Compensating shorter time series with larger numbers of persons and vice versa. *Structural Equation Modeling: A Multidisciplinary Journal*, 28(2), 229–236. <https://doi.org/10.1080/10705511.2020.1779069>
- Hecht, M., Hardt, K., Driver, C. C., & Voelkle, M. C. (2019). Bayesian continuous-time Rasch models. *Psychological Methods*, 24(4), 516–537. <https://doi.org/10.1037/met0000205>
- Hecht, M., Gische, C., Vogel, D., & Zitzmann, S. (2020). Integrating out nuisance parameters for computationally more efficient Bayesian estimation—an illustration and tutorial. *Structural Equation Modeling: A Multidisciplinary Journal*, 27(3), 483–493. <https://doi.org/10.1080/10705511.2019.1647432>
- Hsiao, C. (2014). *Analysis of panel data* (No. 54). Cambridge University Press.
- Jongerling, J., & Hamaker, E. L. (2011). On the trajectories of the predetermined ALT model: What are we really modeling? *Structural Equation Modeling: A Multidisciplinary Journal*, 18(3), 370–382. <https://doi.org/10.1080/10705511.2011.582004>
- Kuiper, R. M., & Ryan, O. (2018). Drawing conclusions from cross-lagged relationships: Re-considering the role of the time-interval. *Structural Equation Modeling: A Multidisciplinary Journal*, 25(5), 809–823. <https://doi.org/10.1080/10705511.2018.1431046>
- Little, T. D. (2013). *Longitudinal structural equation modeling*. Guilford press.

- Loossens, T., Tuerlinckx, F., & Verdonck, S. (2021). A comparison of continuous and discrete time modeling of affective processes in terms of predictive accuracy. *Scientific Reports*, 11(1), 6218. <https://doi.org/10.1038/s41598-021-85320-4>
- Lütkepohl, H. (2005). New introduction to multiple time series analysis. *Springer Science & Business Media*. <https://doi.org/10.1007/978-3-540-27752-1>
- McArdle, J. J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology*, 60, 577–605. <https://doi.org/10.1146/annurev.psych.60.110707.163612>
- McArdle, J. J., & Hamagami, F. (2001). Latent difference score structural models for linear dynamic analyses with incomplete longitudinal data. In L. Collins & A. Sayer (Eds.), *New methods for the analysis of change* (pp. 139–175). American Psychological Association.
- Millsap, R. E. (2008). Introduction to the special issue on growth models for longitudinal data in educational research. *Educational Research and Evaluation*, 14(4), 283–285. <https://doi.org/10.1080/13803610802249308>
- Neale, M. C., Hunter, M. D., Pritikin, J. N., Zahery, M., Brick, T. R., Kirkpatrick, R. M., Estabrook, R., Bates, T. C., Maes, H. H., & Boker, S. M. (2016). OpenMx 2.0: Extended structural equation and statistical modeling. *Psychometrika*, 81(2), 535–549. <https://doi.org/10.1007/s11336-014-9435-8>
- Núñez-Regueiro, F., Juhel, J., Bressoux, P., & Nurra, C. (2021). Identifying reciprocities in school motivation research: A review of issues and solutions associated with cross-lagged effects models. *Journal of Educational Psychology*. <https://doi.org/10.1037/edu0000700>
- OECD. (2009). *PISA data analysis manual SPSS*. OECD Publishing.
- OECD. (2014). *PISA 2012 results: What students know and can do (Volume 1, Revised edition): Student performance in mathematics, reading and science*. Author. <https://doi.org/10.1787/9789264208780-en>
- OECD. (2019b). PISA 2018 assessment and analytical framework. *PISA*, OECD Publishing. <https://doi.org/10.1787/b25efab8-en>
- Oravec, Z., Tuerlinckx, F., & Vandekerckhove, J. (2011). A hierarchical latent stochastic differential equation model for affective dynamics. *Psychological Methods*, 16, 468–490. <https://doi.org/10.1080/00273171.2015.1110512>
- Oravec, Z., Wood, J., & Ram, N. (2018). On fitting a continuous-time stochastic process model in the Bayesian framework. In K. Van Montfort, J. H. L. Oud, & M. C. Voelkle (Eds.), *Continuous time modeling in the behavioral and related sciences* (pp. 55–78). Springer. [https://doi.org/10.1007/978-3-319-77219-6\\_3](https://doi.org/10.1007/978-3-319-77219-6_3)
- Oud, J. H. L. (2010). Second-order stochastic differential equation model as an alternative for the ALT and CALT models. *AStA Advances in Statistical Analysis*, 94(2), 203–215. <https://doi.org/10.1007/s10182-010-0131-4>
- Oud, J. H., & Jansen, R. A. (2000). Continuous time state space modeling of panel data by means of SEM. *Psychometrika*, 65(2), 199–215. <https://doi.org/10.1007/BF02294374>
- Oud, J. H., & Voelkle, M. C. (2014). Do missing values exist? Incomplete data handling in cross-national longitudinal studies by means of continuous time modeling. *Quality & Quantity*, 48(6), 3271–3288. <https://doi.org/10.1007/s11135-013-9955-9>
- Oud, J. H., Voelkle, M. C., & Driver, C. C. (2018). First-and higher-order continuous time models for arbitrary N using SEM. In K. Van Montfort, J. H. L. Oud, & M. C. Voelkle (Eds.), *Continuous time modeling in the behavioral and related sciences* (pp. 1–26). Cham: Springer. [https://doi.org/10.1007/978-3-319-77219-6\\_1](https://doi.org/10.1007/978-3-319-77219-6_1)
- Robitzsch, A., Grund, S., Henke, T., & Robitzsch, M. A. (2017). *Package 'miceadds'*. R Package.
- Rubin, D. B. (2004). *Multiple imputation for nonresponse in surveys* (Vol. 81). Wiley.
- Ryan, O., & Hamaker, E. L. (2021). Time to intervene: A continuous-time approach to network analysis and centrality. *Psychometrika*. <https://doi.org/10.31234/osf.io/2ambn>
- Ryan, O., Kuiper, R. M., & Hamaker, E. L. (2018). A continuous-time approach to intensive longitudinal data: what, why, and how? In K. Van Montfort, J. H. L. Oud, & M. C. Voelkle (Eds.), *Continuous time modeling in the behavioral and related sciences* (pp. 27–54). Springer. [https://doi.org/10.1007/978-3-319-77219-6\\_2](https://doi.org/10.1007/978-3-319-77219-6_2)
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6, 461–464.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton.
- Singer, J. D. (2019). Reshaping the arc of quantitative educational research: It's time to broaden our paradigm. *Journal of Research on Educational Effectiveness*, 12(4), 570–593. <https://doi.org/10.1080/19345747.2019.1658835>
- Sivo, S., & Fan, X. (2008). The latent curve ARMA (p, q) panel model: Longitudinal data analysis in educational research and evaluation. *Educational Research and Evaluation*, 14(4), 363–376. <https://doi.org/10.1080/13803610802249670>
- Specht, J. (Ed.). (2017). *Personality development across the lifespan*. Elsevier Academic Press.
- Usami, S., Murayama, K., & Hamaker, E. L. (2019). A unified framework of longitudinal models to examine reciprocal relations. *Psychological Methods*, 24(5), 637–657. <https://doi.org/10.1037/met0000210>
- Voelkle, M. C. (2008). Reconsidering the use of autoregressive latent trajectory (ALT) models. *Multivariate Behavioral Research*, 43(4), 564–591.
- Voelkle, M. C., & Oud, J. H. (2013). Continuous time modelling with individually varying time intervals for oscillating and non-oscillating processes. *British Journal of Mathematical and Statistical Psychology*, 66(1), 103–126. <https://doi.org/10.1111/j.2044-8317.2012.02043.x>
- Voelkle, M. C., & Oud, J. H. L. (2015). Relating latent change score and continuous time models. *Structural Equation Modeling*, 22(3), 366–381. <https://doi.org/10.1080/10705511.2014.935918>
- Voelkle, M. C., Oud, J. H., Davidov, E., & Schmidt, P. (2012). An SEM approach to continuous time modeling of panel data: Relating authoritarianism and anomia. *Psychological Methods*, 17(2), 176. <https://doi.org/10.1037/a0027543>
- Voelkle, M. C., Gische, C., Driver, C. C., & Lindenberger, U. (2018). The role of time in the quest for understanding psychological mechanisms. *Multivariate Behavioral Research*, 53(6), 782–805. <https://doi.org/10.1080/00273171.2018.1496813>
- Walls, T. A., & Schafer, J. L. (Eds.). (2006). *Models for intensive longitudinal data*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195173444.001.0001>
- Weirich, S., Hecht, M., Becker, B., & Zitzmann, S. (2021). Comparing group means with the total mean in random samples, surveys, and large-scale assessments: A tutorial and software illustration. *Behavior Research Methods*. <https://doi.org/10.3758/s13428-021-01553-1>

- Wigfield, A., Eccles, J. S., Schiefele, U., Roeser, R. W., & Davis-Kean, P. (2006). Development of achievement motivation. In N. Eisenberg, W. Damon, & R. M. Lerner (Eds.), *Handbook of child psychology: Social, emotional, and personality development* (pp. 933–1002). Wiley.
- Wigfield, A., Eccles, J. S., Fredricks, J. A., Simpkins, S., Roeser, R. W., & Schiefele, U. (2015). Development of achievement motivation and engagement. In M. E. Lamb & R. M. Lerner (Eds.), *Handbook of child psychology and developmental science. Socioemotional processes* (7th ed., Vol. 3, pp. 657–700). Wiley. <https://doi.org/10.1002/9781118963418.childpsy316>
- Wu, M. (2005). The role of plausible values in large-scale surveys. *Studies in Educational Evaluation*, 31(2–3), 114–128. <https://doi.org/10.1016/j.stueduc.2005.05.005>
- Zirkel, S., Garcia, J. A., & Murphy, M. C. (2015). Experience-sampling research methods and their potential for education research. *Educational Researcher*, 44(1), 7–16. <https://doi.org/10.3102/0013189X14566879>
- Zyphur, M. J., Allison, P. D., Tay, L., Voelkle, M. C., Preacher, K. J., Zhang, Z., Hamaker, E. L., Shamsollahi, A., Pierides, D. C., Koval, P., & Diener, E. (2020). From data to causes I: Building a general cross-lagged panel model (GCLM). *Organizational Research Methods*, 23(4), 651–687. <https://doi.org/10.1177/1094428119847278>
- Zyphur, M. J., Voelkle, M. C., Tay, L., Allison, P. D., Preacher, K. J., Zhang, Z., Hamaker, E. L., Shamsollahi, A., Pierides, D. C., Koval, P., & Diener, E. (2020). From data to causes II: Comparing approaches to panel data analysis. *Organizational Research Methods*, 23(4), 688–716. <https://doi.org/10.1177/1094428119847280>
- Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In: B. N. Petrov & F. Csáki (Eds.), *2nd international symposium on information theory, Tsahkadsor, Armenia, USSR*, September 2–8, 1971 (pp. 267–281). Budapest: Akadémiai Kiadó. [https://doi.org/10.1007/978-1-4612-1694-0\\_15](https://doi.org/10.1007/978-1-4612-1694-0_15)
- Driver, C. C., & Voelkle, M. C. (2017). Introduction to hierarchical continuous time dynamic modelling with ctsem. R package Vignette. <https://cran.r-project.org/web/packages/ctsem/index.html>
- Driver, C. C. (2020, April 27). Latent growth curves, state dependent error. <https://cdriver.netlify.app/post/lgc/>
- Driver, C. C., & Voelkle, M. C. (2021). Hierarchical continuous time modeling. In *The handbook of personality dynamics and processes* (pp. 887–908). Academic Press. <https://doi.org/10.1016/B978-0-12-813995-0.00034-0>
- Driver, C. C. (2022, May 10). How to model a quadratic trend + independent dynamics. <https://github.com/cdriveraus/ctsem/discussions/29>
- Fisher, M. (2001). Modeling negative autoregression in continuous time. [http://www.markfisher.net/~mefisher/papers/continuous\\_ar.pdf](http://www.markfisher.net/~mefisher/papers/continuous_ar.pdf)
- Jindra, C., Hardt, K., & Hecht, M. (under review). Dynamics between reading and math proficiency over time in secondary education—observational evidence from continuous time models [Submitted for publication].
- Lüdtke, O., & Robitzsch, A. (2021, July 29). A critique of the random intercept cross-lagged panel model. <https://doi.org/10.31234/osf.io/6f85c>
- OECD (2019a), "Which countries have improved and which countries have declined in performance over their participation in PISA?". In *PISA 2018 results (volume I): What students know and can do*, OECD Publishing. <https://doi.org/10.1787/cf74dac9-en>
- OECD (2021, January 13). PISA database. <https://www.oecd.org/pisa/data/>
- Oud, J. H., & Delsing, M. J. (2010). Continuous time modeling of panel data by means of SEM. In *Longitudinal research with latent variables* (pp. 201–244). Springer. [https://doi.org/10.1007/978-3-642-11760-2\\_7](https://doi.org/10.1007/978-3-642-11760-2_7)
- R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Stan Development Team (2020). RStan: the R interface to Stan. R package version 2.21.2. <http://mc-stan.org/>.
- van Buuren, S., Groothuis-Oudshoorn, K., Robitzsch, A., Vink, G., Doove, L., & Jolani, S. (2015). Package 'mice'. *Computer software*.
- van Montfort, K., Oud, J. H. L., & Voelkle, M. C. (Eds.). (2018). *Continuous time modeling in the behavioral and related sciences*. Springer.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)